```
                GenCore version 5.1.3
       Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         January 27, 2003, 16:41:03 ; Search time 82 Seconds
                                             (without alignments)
                                             11982.833 Million cell updates/sec

Title:          US-09-830-433A-7
Perfect score:  3204
Sequence:       1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	49.6	1.5	913	4	US-08-818-112-16	Sequence 16, Appl
	2	49.6	1.5	913	4	US-08-818-111-16	Sequence 16, Appl
	3	49.6	1.5	913	4	US-09-056-556-16	Sequence 16, Appl
	4	49.6	1.5	913	4	US-09-072-596-16	Sequence 16, Appl
	5	49.6	1.5	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	6	49.6	1.5	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	7	47.8	1.5	7218	1	US-08-232-463-14	Sequence 14, Appl
c	8	45.2	1.4	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	9	45.2	1.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	10	41.4	1.3	2949	4	US-09-412-554A-3	Sequence 3, Appli
	11	39.4	1.2	2367	4	US-09-056-556-201	Sequence 201, App
	12	39.4	1.2	2367	4	US-09-072-596-196	Sequence 196, App
	13	39.2	1.2	625	4	US-08-998-416-1112	Sequence 1112, Ap
c	14	37.8	1.2	329	4	US-09-056-556-168	Sequence 168, App
c	15	37.8	1.2	329	4	US-09-072-596-163	Sequence 163, App
c	16	37.6	1.2	494	4	US-09-056-556-176	Sequence 176, App
c	17	37.6	1.2	494	4	US-09-072-596-171	Sequence 171, App
c	18	37.4	1.2	3572	4	US-09-575-574-3	Sequence 3, Appli
	19	37	1.2	610	4	US-09-072-596-293	Sequence 293, App
c	20	36.8	1.1	954	2	US-08-336-198C-2	Sequence 2, Appli
c	21	36.6	1.1	6530	2	US-08-146-930-1	Sequence 1, Appli
c	22	36.6	1.1	6530	3	US-08-458-240-1	Sequence 1, Appli
c	23	36.6	1.1	6530	5	PCT-US93-03993-1	Sequence 1, Appli
	24	36.4	1.1	1631	4	US-08-895-590-43	Sequence 43, Appl
c	25	36.2	1.1	1280	4	US-09-060-756-4	Sequence 4, Appli
	26	36.2	1.1	1399	1	US-08-471-033-24	Sequence 24, Appl
	27	36.2	1.1	1399	2	US-08-471-044-24	Sequence 24, Appl
	28	36.2	1.1	1399	2	US-08-463-483A-24	Sequence 24, Appl
	29	36.2	1.1	1399	2	US-08-471-046A-24	Sequence 24, Appl
	30	36.2	1.1	1399	2	US-08-470-566B-24	Sequence 24, Appl
	31	36.2	1.1	1399	2	US-08-469-334-24	Sequence 24, Appl
	32	36.2	1.1	1399	3	US-09-300-529-24	Sequence 24, Appl
c	33	36	1.1	935	1	US-08-162-475A-3	Sequence 3, Appli
	34	35.6	1.1	4162	2	US-08-459-448A-26	Sequence 26, Appl
	35	35.6	1.1	4162	3	US-08-459-595A-26	Sequence 26, Appl
	36	35.6	1.1	4162	3	US-08-459-504B-26	Sequence 26, Appl
	37	35.6	1.1	4162	3	US-08-459-444-26	Sequence 26, Appl
	38	35.6	1.1	4162	4	US-09-547-422-26	Sequence 26, Appl
	39	35.6	1.1	4165	1	US-07-951-715A-26	Sequence 26, Appl
	40	35.6	1.1	4689	4	US-09-105-537-34	Sequence 34, Appl
	41	35.6	1.1	36778	4	US-09-105-537-5	Sequence 5, Appli
	42	35.6	1.1	38506	3	US-09-320-878-19	Sequence 19, Appl
	43	35.2	1.1	538	4	US-09-056-556-180	Sequence 180, App
	44	35.2	1.1	538	4	US-09-072-596-175	Sequence 175, App
c	45	35.2	1.1	991	4	US-09-344-529-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-08-818-112-16
; Sequence 16, Application US/08818112
; Patent No. 6290969
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  153
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,112
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  913 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-112-16

  Query Match             1.5%;  Score 49.6;  DB 4;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.00048;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;
```

```
Qy    2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
            ||| |  |   || |||||||||        ||||| | | |  || |    ||
Db     605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy    2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |        || |  |  |  ||       ||||     |||   |  |||
Db     665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy    2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
            |||| |  ||||   || | | |     |:||   | ||      || |     |   | ||
Db     725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy    2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
            | || |  |||||||||    |     |  | || || ||     |||    |  || |  |
Db     785 CCCGCCGTTGCCGTACAGCCACCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy    2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
            |||| | ||| |   ||||| |  | ||||||||   ||||   ||
Db     845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889
```

RESULT 2
US-08-818-111-16
; Sequence 16, Application US/08818111
; Patent No. 6338852
; GENERAL INFORMATION:
;   APPLICANT: Reed, Steven G.
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Dillon, Davin C.
;   APPLICANT: Campos-Neto, Antonia
;   APPLICANT: Houghton, Raymond
;   APPLICANT: Vedvick, Thomas S.
;   APPLICANT: Twardzik, Daniel R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF   TUBERCULOSIS
;   NUMBER OF SEQUENCES: 148
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEED and BERRY LLP
;   STREET: 6300 Columbia Center, 701 Fifth Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: USA
;   ZIP: 98104-7092
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/818,111
;   FILING DATE: 13-MAR-1997
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Maki, David J.
;   REGISTRATION NUMBER: 31,392
;   REFERENCE/DOCKET NUMBER: 210121.417C6
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 16:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 913 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-818-111-16

```
  Query Match             1.5%;  Score 49.6;  DB 4;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.00048;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy    2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
            ||| |  |   || |||||||||        ||||| | | |  || |    ||
Db     605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy    2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |        || |  |  |  ||       ||||     |||   |  |||
Db     665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy    2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
            |||| |  ||||   || | | |     |:||   | ||      || |     |   | ||
Db     725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy    2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
            | || |  |||||||||    |     |  | || || ||     |||    |  || |  |
Db     785 CCCGCCGTTGCCGTACAGCCACCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy    2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
            |||| | ||| |   ||||| |  | ||||||||   ||||   ||
Db     845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889
```

RESULT 3
US-09-056-556-16
; Sequence 16, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
;   APPLICANT: Reed, Steven G.
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Dillon, Davin C.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND   TRE
;   NUMBER OF SEQUENCES: 241
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEED and BERRY LLP
;   STREET: 6300 Columbia Center, 701 Fifth Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: USA
;   ZIP: 98104-7092
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/056,556
;   FILING DATE: 07-APR-1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Maki, David J.
;   REGISTRATION NUMBER: 31,392
;   REFERENCE/DOCKET NUMBER: 210121.457
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 16:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 913 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-09-056-556-16

```
  Query Match             1.5%;  Score 49.6;  DB 4;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.00048;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy    2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
            ||| |  |   || |||||||||        ||||| | | |  || |    ||
Db     605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy    2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |        || |  |  |  ||       ||||     |||   |  |||
Db     665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy    2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
            |||| |  ||||   || | | |     |:||   | ||      || |     |   | ||
```

```
Db     725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy    2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
              | || |  |||||||||    |      |   | || || ||    |||   |  || |  |
Db     785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy    2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
             |||| | ||| |  ||||| |  | ||||||||   ||||   ||
Db     845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 4
US-09-072-596-16
; Sequence 16, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  913 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-16

  Query Match             1.5%;  Score 49.6;  DB 4;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.00048;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy    2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
             ||| |   |    || |||||||||            |||||  | | |   || |    ||
Db     605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy    2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                   |         || |  |   |   ||        ||||      |||   |  |||
Db     665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy    2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
              |||| |  ||||   || | | |     |:||   | ||      || |     |   | ||
Db     725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784
```

```
Qy    2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
              | || |  |||||||||    |      |   | || || ||    |||   |  || |  |
Db     785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy    2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
             |||| | ||| |  ||||| |  | ||||||||   ||||   ||
Db     845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 5
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             1.5%;  Score 49.6;  DB 4;  Length 4403765;
  Best Local Similarity   44.5%;  Pred. No. 0.041;
  Matches  196;  Conservative    0;  Mismatches  244;  Indels    0;  Gaps    0;

Qy    2076 CGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGAT 2135
             |||  |||||| |          |||||  | |    |  | ||    || |    ||     | |
Db  336988 CGCTGCCGCCGCTGCCGCCGGCGCCGCCGGAGGCGCCGAGGGCGAGTAGGCCGGCGTTGC 337047

Qy    2136 GGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGC 2195
              | ||      |   |     |||     |     ||    |   | |     | |      || ||      ||
Db  337048 CGCCGCTGCCGCCGTTGCCGCCGAAGGTGCCGCCGCTGCTGCCGCCAGCACCGCCAGTGC 337107

Qy    2196 CGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGC 2255
             || |  | ||  |  | ||     ||| |  |  |  |  | ||||  || |||          |   ||        ||
Db  337108 CGCCGGCGCCGAACAGCCCGCCGTGCCCCCCGGCGCCGCCGTCGGCGCCGAGCGTGCCCG 337167

Qy    2256 CGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGT 2315
             |           |||       | ||||||  |       |    |      |     |   | ||| ||||    |   ||
Db  337168 CCCCGCCGGTGCCGCCGGCGCCGAAGAGCAATCCGTTCCCCCCGGTCCCGCCATTCGCGC 337227

Qy    2316 CTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATC 2375
             | |   ||| |  |   |||||| ||  |||| |      |        | |        |     ||| ||
Db  337228 CAAACCCGCCGGCCCCGCCGGCCCCGCCGTTGGCGAACAGCCCACCGGTACCACCGGCTC 337287

Qy    2376 GGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGG 2435
              | |||     |   |   || || ||   |  |  || |   |        ||    |     ||           | |
Db  337288 CGGCGGTGCCGCCGGCACCGATAAAGTTTGGGAGAGGGCGGCCTGGCCGCCGGTCCCTG 337347

Qy    2436 TGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGG 2495
              || ||    |||||| | ||| ||       | |             ||||      ||        |     |
Db  337348 CGGCACCGAGGAACAAGCCGGCGTCACCGCCGCGCCCGCCGGCCCCGCCGGTGTCCAGGC 337407

Qy    2496 CATTGCCGCGAAAACCGGCG 2515
             ||    ||||       ||| ||
Db  337408 CAAACCCGCCGCTGCCGCCG 337427
```

```
RESULT 6
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             1.5%;  Score 49.6;  DB 4;  Length 4411529;
  Best Local Similarity   44.1%;  Pred. No. 0.041;
  Matches  208;  Conservative    0;  Mismatches  264;  Indels    0;  Gaps    0;

Qy    2044 GCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTA 2103
           || |  |   | | | | || |       |    |||  |||||| |        |||||
Db 336867 GCCGACCCCGCCGGTCCCGCCGACATTTAAGGCGCTGCCGCCGCTGCCGCCGGCGCCGCC 336926

Qy    2104 GAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC 2163
           | |   | ||   || |   ||    | |    | ||       |  |    |||    |
Db 336927 GGAGGCGCCGAGGGCGAGTAGGCCGGCGTTGCCGCCGCTGCCGCCGTTGCCGCCGAAGGT 336986

Qy    2164 GCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGC 2223
           || |  | |   | |      || ||    |||| | || | | ||   ||| | | | | |
Db 336987 GCCGCCGCTGCTGCCGCCAGCACCGCCAGTGCCGCCGGCGCCGAACAGCCCGCCGTGCCC 337046

Qy    2224 CCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGT 2283
           |||  || |||      |  ||        ||    |           |||     | |||||| |
Db 337047 CCCGGCGCCGCCGTCGGCGCCGAGCGTGCCCGCCCCGCCGGTGCCGCCGGCGCCGAAGAG 337106

Qy    2284 GTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCC 2343
              |  |    |  | ||| ||||  |   | |  |   ||| |   | |||||| ||  |||
Db 337107 CAATCCGTTCCCCCCGGTCCCGCCATTCGCGCCAAACCCGCCGGCCCCGCCGGCCCCGCC 337166

Qy    2344 GATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGT 2403
           | |     |     | |       |    |||  || | |||      | |  || || ||  | |
Db 337167 GTTGGCGAACAGCCCACCGGTACCACCGGCTCCGGCGGTGCCGCCGGCACCGATAAAGTT 337226

Qy    2404 CTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAA 2463
           || | |     ||   |    ||              | | || ||    |||||| | ||| ||
Db 337227 TTGGGAGAGGGCGGCCTGGCCGCCGGTCCCTGCGGCACCGAGGAACAAGCCGGCGTCACC 337286

Qy    2464 GGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCG 2515
           | |          ||||     ||       |    | ||     ||||       ||| ||
Db 337287 GCCGCGCCCGCCGGCCCCGCCGGTGTCCAGGCCAAACCCGCCGCTGCCGCCG 337338


RESULT 7
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             1.5%;  Score 47.8;  DB 1;  Length 7218;
  Best Local Similarity   6.0%;  Pred. No. 0.0045;
  Matches   25;  Conservative  216;  Mismatches  178;  Indels    0;  Gaps    0;

Qy     432 CATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAA 491
           | |||| | |  | | | |  | |  | | |||||   ||    :: :::::  : :::::
Db    1476 CCTATCTATGCAAGTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRR 1417

Qy     492 CTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGC 551
              : :  :: : : : : ::::::::: :    ::::: ::::: :: ::::: :   ::::
Db    1416 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1357

Qy     552 TTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGT 611
                   :: :: ::::  :   : :::::  :::: ::::  :: ::: :    :  :  :
Db    1356 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1297

Qy     612 AAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGG 671
           :::::::  ::: : :    :: : ::         : :   :: ::: :      : ::: ::
Db    1296 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1237

Qy     672 CAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGG 731
              :::   : ::: :: :    :  :    :: : :: :    : : : :: ::: : : : :: ::
Db    1236 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1177

Qy     732 AACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACG 791
            ::  ::::: :::: :: :    : ::   :    : :: : : :::   :::   ::: ::: :
Db    1176 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1117

Qy     792 TGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACC 850
             :: : : : :  :  :: :: ::    :::: :: :  ::::: ::: : |  | | |
Db    1116 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGC 1058


RESULT 8
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
```

```
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             1.4%;  Score 45.2;  DB 4;  Length 4403765;
  Best Local Similarity   57.3%;  Pred. No. 0.68;
  Matches  102;  Conservative    0;  Mismatches   73;  Indels    3;  Gaps    1;

Qy    1225 GGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATAC 1284
           || ||   |||||| ||  ||  ||| | ||||   ||||| ||| ||| |    |||| |
Db 4361104 GGCACCAGCTTTTCGGCGGCCTACGTCAGCGGCGTCGCGGCACTGGTGCGGGCCAAATTC 4361045

Qy    1285 CCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGT 1344
           ||   | |||    | ||| |  ||    | |  | |  ||     |||| ||  |         |
Db 4361044 CCCGAGTTGACTGCCTACCAAGTGATCAACCGGATCGTGCAGTCGGCGCACAATCCGCCT 4360985

Qy    1345 GCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGA 1402
           ||    ||||||||| |||  | || |  || ||| ||||| |||      ||   |||
Db 4360984 GCG---GGCGTGGACAACAAACTGGGGTACGGGCTGGTGGATCCGGTCGCCGCATTGA 4360930

RESULT 9
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             1.4%;  Score 45.2;  DB 4;  Length 4411529;
  Best Local Similarity   57.3%;  Pred. No. 0.68;
  Matches  102;  Conservative    0;  Mismatches   73;  Indels    3;  Gaps    1;

Qy    1225 GGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATAC 1284
           || ||   |||||| ||  ||  ||| | ||||   ||||| ||| ||| |    |||| |
Db 4368871 GGCACCAGCTTTTCGGCGGCCTACGTCAGCGGCGTCGCGGCACTGGTGCGGGCCAAATTC 4368812

Qy    1285 CCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGT 1344
           ||   | |||    | ||| |  ||    | |  | |  ||     |||| ||  |         |
Db 4368811 CCCGAGTTGACTGCCTACCAAGTGATCAACCGGATCGTGCAGTCGGCGCACAATCCGCCT 4368752

Qy    1345 GCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGA 1402
           ||    ||||||||| |||  | || |  || ||| ||||| |||      ||   |||
Db 4368751 GCG---GGCGTGGACAACAAACTGGGGTACGGGCTGGTGGATCCGGTCGCCGCATTGA 4368697

RESULT 10
US-09-412-554A-3
; Sequence 3, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
;  APPLICANT:  Conklin, Darrell
;  APPLICANT:  Ellsworth, Jeff L.
;  TITLE OF INVENTION:  FOLLISTATIN RELATED PROTEIN ZFSTA2
;  FILE REFERENCE:  98-50
;  CURRENT APPLICATION NUMBER: US/09/412,554A
;  CURRENT FILING DATE:  1999-10-05
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 2949
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Degenerate oligonucleotide sequence encoding the
;   OTHER INFORMATION: zfsta2 polypeptide of SEQ ID NO:2.
;   NAME/KEY: variation
;   LOCATION: (1)...(2949)
;   OTHER INFORMATION: Each N is independently any nucleotide.
US-09-412-554A-3

  Query Match             1.3%;  Score 41.4;  DB 4;  Length 2949;
  Best Local Similarity   32.5%;  Pred. No. 0.17;
  Matches   63;  Conservative   33;  Mismatches   98;  Indels    0;  Gaps    0;

Qy     452 ATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGA 511
           |:|| ::  |:|  ||:  ::  ||:|  :  ||: |: |:         || ||:|   | |
Db    1589 AYGGNWSNGARGTNCAYATHWSNAAYGTNMGNTAYGARGAYACNGGNGCNTAYACNTGYA 1648

Qy     512 AGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCG 571
           |    :  :||:|  || |  | : |:||:||::   : :  ||:|  || ||  : |
Db    1649 THGCNAARAAYGARGCNGGNGTNGAYGARGAYATHWSNWSNYTNTTYGTNGARGAYWSNG 1708

Qy     572 TTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCG 631
             :   |:||    || ||            :          | |     |  ||:||  |:||
Db    1709 CNMGNAARACNYTNGCNAAYATHYTNTGGMGNGARGARGGNYTNGGNATHGGNAAYATGT 1768

Qy     632 ATGTGGTCTCCCAT 645
              :   || |  : |:
Db    1769 TYTAYGTNTTYTAY 1782

RESULT 11
US-09-056-556-201
; Sequence 201, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND          TRE
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  201:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2367 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-201

  Query Match             1.2%;  Score 39.4;  DB 4;  Length 2367;
  Best Local Similarity   50.8%;  Pred. No. 0.54;
  Matches   94;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy    2164 GCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGC 2223
              || ||     |    | ||||    | |    |||||  || |  |  ||      | |||   |   |
Db     119 GCCACCGGAGCCGCCGTTGCCGCCGTCACCGCCAGCCCCGCCGACTCCACCGGCCCCACC 178

Qy    2224 CCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGT 2283
               || || |||  |    |    || || |  ||||  || |||||   || ||||   |   |
Db     179 GACTCCGCCGCTGCCACCGTTGCCGCCGTTGCCGATCAACATGCCGCTGGCGCCACCCTT 238

Qy    2284 GTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCC 2343
             |     |   |  || |  |    ||||     ||| |  |   || |    |    |  ||||||    |||
Db     239 GCCACCCACGCCACCGGCTCCGCCCACCCCGCCGACACCAAGCGAGCTGCCGCCGGAGCC 298

Qy    2344 GATAT 2348
              ||
Db     299 ACCAT 303


RESULT 12
US-09-072-596-196
; Sequence 196, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2367 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-196

  Query Match             1.2%;  Score 39.4;  DB 4;  Length 2367;
  Best Local Similarity   50.8%;  Pred. No. 0.54;
  Matches   94;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy    2164 GCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGC 2223
              || ||     |    | ||||    | |    |||||  || |  |  ||      | |||   |   |
Db     119 GCCACCGGAGCCGCCGTTGCCGCCGTCACCGCCAGCCCCGCCGACTCCACCGGCCCCACC 178

Qy    2224 CCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGT 2283
               || || |||  |    |    || || |  ||||  || |||||   || ||||   |   |
Db     179 GACTCCGCCGCTGCCACCGTTGCCGCCGTTGCCGATCAACATGCCGCTGGCGCCACCCTT 238

Qy    2284 GTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCC 2343
             |     |   |  || |  |    ||||     ||| |  |   || |    |    |  ||||||    |||
Db     239 GCCACCCACGCCACCGGCTCCGCCCACCCCGCCGACACCAAGCGAGCTGCCGCCGGAGCC 298

Qy    2344 GATAT 2348
              ||
Db     299 ACCAT 303


RESULT 13
US-08-998-416-1112
; Sequence 1112, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
```

```
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  1112:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  625 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1677UP
US-08-998-416-1112

  Query Match             1.2%;  Score 39.2;  DB 4;  Length 625;
  Best Local Similarity   48.6%;  Pred. No. 0.31;
  Matches  107;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy    2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
             | | ||||  |  |||        | |  ||  | ||    |  || |||| ||| |   ||
Db     290 GCCGGCACCTACGTGCGCTCGCCACTCAACTGGGCCGTCTCCGTCGGCGCCGCGTCGCCC 349

Qy    2257 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTC 2316
              |  |||| | | |     ||||| |  ||| ||  |      | || ||  | || |
Db     350 CTGGAGCACGTGGACCAGCTGGACGGCGGCCGCGTCGGCGTGTCACGCGTCGGCAGCGGC 409

Qy    2317 TATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCG 2376
             |    || ||  | |    |  |  |||      || ||    ||| | |       |  ||
Db     410 TCGTACGTCATGAGCTATGTGCTCGCCCTGCAGCGCGGCTTCCGCCGGCCCTTTGCCGCG 469

Qy    2377 GACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCG 2416
              |   | ||  ||||| |  | | | || ||||||| |  ||
Db     470 CATCCGGTGTGCCACACCTTTGCCGGCCTGCGCGCCGCCG 509


RESULT 14
US-09-056-556-168/c
; Sequence 168, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND          TREATM
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  168:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  329 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-168

  Query Match             1.2%;  Score 37.8;  DB 4;  Length 329;
  Best Local Similarity   49.3%;  Pred. No. 0.54;
  Matches   99;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy    2150 CCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATA 2209
            || |||   | | ||    |  |  |   |  |   | || | | ||| |  ||| |  |
Db     244 CCGCCGGCACCTGCGTTGAAGGCCTGGTTGCCGGGGCCGCCGGCCCCGCCGTCACCGCCA 185

Qy    2210 TGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGA 2269
               | | |  || | ||      |||        |  ||  ||     ||||  ||||| ||
Db     184 ACACCGCCGGCCCCGCCGGTCCCGCCGGCCCCGCCGGCGCCTTGGCCGCCCAGCAGGCTG 125

Qy    2270 ATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTA 2329
            ||   |||| |||       |       |  ||| | ||||      ||| |  || |||  |  ||
Db     124 ATCAGGCCGCCGGCCCCGCCGGGGCCGCCAGCCCCGCCAGCCCCGCCCATCCCGCCGTTA 65

Qy    2330 CCGCCGCCCATGCCGATATGC 2350
            || ||  |    |||| ||| |
Db      64 CCACCATCACCGCCGTTATCC 44


RESULT 15
US-09-072-596-163/c
; Sequence 163, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  163:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  329 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-163

  Query Match             1.2%;  Score 37.8;  DB 4;  Length 329;
```

```
  Best Local Similarity   49.3%;  Pred. No. 0.54;
  Matches   99;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy     2150 CCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATA 2209
              || |||    | | ||     |   |   |     |   |     | || | | ||| |   ||| |  |
Db      244 CCGCCGGCACCTGCGTTGAAGGCCTGGTTGCCGGGGCCGCCGGCCCCGCCGTCACCGCCA 185

Qy     2210 TGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGA 2269
                 | | |   || | ||     |||         |   ||   ||     ||||   ||||| ||
Db      184 ACACCGCCGGCCCCGCCGGTCCCGCCGGCCCCGCCGGCGCCTTGGCCGCCCAGCAGGCTG 125

Qy     2270 ATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTA 2329
              ||    |||| |||         |         |   ||| | ||||     ||| | || ||| |  ||
Db      124 ATCAGGCCGCCGGCCCCGCCGGGGCCGCCAGCCCCGCCAGCCCCGCCCATCCCGCCGTTA 65

Qy     2330 CCGCCGCCCATGCCGATATGC 2350
              || ||   |    |||| ||| |
Db       64 CCACCATCACCGCCGTTATCC 44


RESULT 16
US-09-056-556-176/c
; Sequence 176, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND          TREATM
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  176:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-176

  Query Match              1.2%;  Score 37.6;  DB 4;  Length 494;
  Best Local Similarity   51.2%;  Pred. No. 0.75;
  Matches   88;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy     2173 GAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGC 2232
              ||  | || |    |  ||||||||||  | | | |       | | || | ||| |    | |
Db      256 GACGCCGTCGCCGCCCCGGCCGCCGACACCGCCGTTGCCGCCGCCACCGGCCACGGTGCC 197

Qy     2233 GCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATC 2292
              |     |  |  ||| |  |  ||||  | |    ||     |||||||| |   |    |     |
Db      196 GACGGTACCGCCGCCGTTGGGGCCGCCCTGACCGCCGTCGCCGCCGAAGCCGCCCTTGCC 137

Qy     2293 TTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCG 2344
               | | |      ||| ||   ||| |    |||| |      ||||| ||     || |
Db      136 GCCGAAAAAGCCGCTGCCGCCGGCGCCGCCGGCGCCGCCGCCACCGCCGCTG 85


RESULT 17
US-09-072-596-171/c
; Sequence 171, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  171:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  494 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-171

  Query Match              1.2%;  Score 37.6;  DB 4;  Length 494;
  Best Local Similarity   51.2%;  Pred. No. 0.75;
  Matches   88;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy     2173 GAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGC 2232
              ||  | || |    |  ||||||||||  | | | |       | | || | ||| |    | |
Db      256 GACGCCGTCGCCGCCCCGGCCGCCGACACCGCCGTTGCCGCCGCCACCGGCCACGGTGCC 197

Qy     2233 GCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATC 2292
              |     |  |  ||| |  |  ||||  | |    ||     |||||||| |   |    |     |
Db      196 GACGGTACCGCCGCCGTTGGGGCCGCCCTGACCGCCGTCGCCGCCGAAGCCGCCCTTGCC 137

Qy     2293 TTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCG 2344
               | | |      ||| ||   ||| |    |||| |      ||||| ||     || |
Db      136 GCCGAAAAAGCCGCTGCCGCCGGCGCCGCCGGCGCCGCCGCCACCGCCGCTG 85


RESULT 18
US-09-575-574-3/c
; Sequence 3, Application US/09575574
; Patent No. 6376750
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Yu, Su-May
;  APPLICANT:  Chao, Yu-Chan
;  TITLE OF INVENTION:  PLANT SEEDLING AND EMBRYO PROMOTER
;  FILE REFERENCE:  08919-047001
;  CURRENT APPLICATION NUMBER: US/09/575,574
;  CURRENT FILING DATE:  2000-05-22
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 3572
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-09-575-574-3

  Query Match             1.2%;  Score 37.4;  DB 4;  Length 3572;
  Best Local Similarity   48.0%;  Pred. No. 2.4;
  Matches  107;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy    2149 GCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGAT 2208
Db    2535 GCCGCCGTAGCCGCCGCCACGGCTGCCACCGTAGCCGCCGCCGCCGCCGCCACGGCCGCC 2476

Qy    2209 ATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCG 2268
Db    2475 GCCGTAGCCGCCGCCGCCACCGTAGCCGCCCTCACGGCGCCCGTAGCCGCCGCCGCCACC 2416

Qy    2269 AATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGT 2328
Db    2415 TCCTCCGTAGCCGCCGCCTCCACGACCGCCGCCGTAGCCGCCACCGCCGCCGCCGTAGCC 2356

Qy    2329 ACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCG 2371
Db    2355 CCCGCCTCCGCCGCCGGAGCGGCGGGACTGGGCCTCATTGACG 2313


RESULT 19
US-09-072-596-293
; Sequence 293, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  293:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  610 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-072-596-293

  Query Match             1.2%;  Score 37;  DB 4;  Length 610;
  Best Local Similarity   49.2%;  Pred. No. 1.2;
  Matches   97;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy    2180 TTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTT 2239
Db      30 TTGCCTCCGGTGCCGCCGACGCCGGCGTTGCCGCCGTTGCCGCCGTTGCCGCCGGTACCG 89

Qy    2240 TCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACA 2299
Db      90 GGGTTTCCTACGGTGCCGCCGCCCGGCAGCATGGCCCCGCTGTTTAGGCCGTTTTCGCCG 149

Qy    2300 GTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCC 2359
Db     150 GCCCCGCCGTCACCGGCTTTGCCGCCATCGCCGCCGTTGCCGCCGCTGGTGGGGTGGCG 209

Qy    2360 GGCTGAAAGCCGTATCG 2376
Db     210 GCCTGGTTGACGTATTG 226


RESULT 20
US-08-336-198C-2/c
; Sequence 2, Application US/08336198C
; Patent No. 5866382
;  GENERAL INFORMATION:
;    APPLICANT:  Hallborn, Johan
;    APPLICANT:  Penttila, Merja
;    APPLICANT:  Ojamo, Heikki
;    APPLICANT:  Keranen, Sirkka
;    APPLICANT:  Hahn-Hagerdal, Barbei
;    APPLICANT:  Waldfridsson, Mats
;    APPLICANT:  Airaksinen, Ulla
;    TITLE OF INVENTION:  XYLOSE UTILIZATION BY RECOMBINANT YEASTS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  301 N. Washington St.
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-3487
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/336,198C
;      FILING DATE:  03-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-205-8000
;      TELEFAX:  703-205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  954 base pairs
;      TYPE:  nucleic acid
```

```
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Pichia stipitis
;       STRAIN:  CBS-6054
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..954
;       OTHER INFORMATION:  /standard_name= "xylose reductase"
;     PUBLICATION INFORMATION:
;       DOCUMENT NUMBER:  FI 901771
;       FILING DATE:  06-APR-1990
US-08-336-198C-2

  Query Match              1.1%;  Score 36.8;  DB 2;  Length 954;
  Best Local Similarity   50.6%;  Pred. No. 1.8;
  Matches   89;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy    2640 CTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGG 2699
              ||   |||| |||||  |   |  |||   | || ||     |  |  ||   | |  |||
Db     193 CTTCGTCAATGGCCTTCTTGACACCGGCACCAACTAACTTTTCGTTGGCGTAATCTTCGG 134

Qy    2700 TGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGG 2759
               | | ||||||  | | |  | |   |   |   ||||| |     | | |   |  ||   |
Db     133 CACCGTCGAACAATCTGTAACCGGTCTTGATAGCACGGTAGATCTGTTCAGAACAGGTGT 74

Qy    2760 CGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGC 2815
              ||  ||| || ||||     |||| ||| |   |   | |    |||| |   |  |  |  |
Db      73 CGACGTCGACTTTCCAACAGCCGAAACCGACGGCTGGCATGTCGTAACCAGAGTTC 18


RESULT 21
US-08-146-930-1/c
; Sequence 1, Application US/08146930
; Patent No. 5958764
;  GENERAL INFORMATION:
;     APPLICANT:  Roop, Dennis R.
;     APPLICANT:  Rothnagel, Joseph A.
;     APPLICANT:  Greenhalgh, David A.
;     TITLE OF INVENTION:  SPECIFIC EXPRESSION VECTORS
;     TITLE OF INVENTION:  AND METHODS OF USE
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  LYON & LYON
;       STREET:  611 West Sixth Street
;       CITY:  Los Angeles
;       STATE:  California
;       COUNTRY:  U.S.A.
;       ZIP:  90017
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  IBM MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/146,930
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;     PRIOR APPLICATION DATA:  including application
;     PRIOR APPLICATION DATA:  described below:             two
;       APPLICATION NUMBER:  07/876,286
;       FILING DATE:  April 30, 1992
;       APPLICATION NUMBER:  No. 5958764 yet assigned (204/132)
;       FILING DATE:  October 29, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warburg, Richard J.
;       REGISTRATION NUMBER:  32,327
;       REFERENCE/DOCKET NUMBER:  204/152
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (213) 489-1600
;       TELEFAX:  (213) 955-0440
;       TELEX:  67-3510
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6530 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-146-930-1

  Query Match              1.1%;  Score 36.6;  DB 2;  Length 6530;
  Best Local Similarity   45.8%;  Pred. No. 5.5;
  Matches  126;  Conservative    0;  Mismatches  149;  Indels    0;  Gaps    0;

Qy    2063 CAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATC 2122
              | |  | || |  ||| ||| || |          ||||      |    || | || |
Db    3142 CCGGAGTACTTGACGCCCCCACCAGAGCCGCCACCGCCTCCACAGCTGGAGCCACCCCCG 3083

Qy    2123 TGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTG 2182
               | |         | || |    |  ||| |||    | ||||  |   |  ||     | |  |
Db    3082 GAGTAGCTGCCGCCGCCGCAGCTGGAGCCGCCGCCGCCTCCGGAGTAGCCACCTCCGCAG 3023

Qy    2183 AAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCC 2242
                   || ||||||| || || ||          | | |             | | |  | | |
Db    3022 CTGGAGCCGCCGCCGCCCCAGAGTAGCCACCGCCGCAGCTAGAGCCACCACCGCCTCCG 2963

Qy    2243 GCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTC 2302
              | | ||| || |||| |          || |  ||||||  ||| ||||      |   |  ||
Db    2962 GAGTAGCCGCCGCCGCAGCTAGAGCCACCGCCGCCTCCGGAGTACTTGACGCCCCCACCG 2903

Qy    2303 TCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
               |||||  ||   |     || |    | ||| || ||
Db    2902 GAGCCGCCTCCACCGCCGCAGCTGGAACCACCTCC 2868


RESULT 22
US-08-458-240-1/c
; Sequence 1, Application US/08458240
; Patent No. 6143727
;  GENERAL INFORMATION:
;     APPLICANT:  Roop, Dennis R.
;     APPLICANT:  Rothnagel, Joseph A.
;     APPLICANT:  Greenhalgh, David A.
;     TITLE OF INVENTION:  SPECIFIC EXPRESSION VECTORS
;     TITLE OF INVENTION:  AND METHODS OF USE
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  LYON & LYON
;       STREET:  611 West Sixth Street
;       CITY:  Los Angeles
;       STATE:  California
;       COUNTRY:  U.S.A.
;       ZIP:  90017
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  IBM MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/458,240
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/146,930
;       FILING DATE:
;       APPLICATION NUMBER:  07/876,286
;       FILING DATE:  April 30, 1992
;       APPLICATION NUMBER:  No. 6143727 yet assigned (204/132)
;       FILING DATE:  October 29, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warburg, Richard J.
;       REGISTRATION NUMBER:  32,327
```

```
;      REFERENCE/DOCKET NUMBER:  204/152
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6530 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-458-240-1

  Query Match              1.1%;  Score 36.6;  DB 3;  Length 6530;
  Best Local Similarity   45.8%;  Pred. No. 5.5;
  Matches  126;  Conservative    0;  Mismatches  149;  Indels    0;  Gaps    0;

Qy   2063 CAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATC 2122
          | |  | || |  ||| ||| || |         ||||     |   || | || |
Db   3142 CCGGAGTACTTGACGCCCCCACCAGAGCCGCCACCGCCTCCACAGCTGGAGCCACCCCCG 3083

Qy   2123 TGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTG 2182
            | |          | || |   |  ||| |||   | |||| |  | ||    | |  |
Db   3082 GAGTAGCTGCCGCCGCCGCAGCTGGAGCCGCCGCCGCCTCCGGAGTAGCCACCTCCGCAG 3023

Qy   2183 AAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCC 2242
               || ||||||| || || ||        | | |          |  | |  | | |
Db   3022 CTGGAGCCGCCGCCGCCCCCAGAGTAGCCACCGCCGCAGCTAGAGCCACCACCGCCTCCG 2963

Qy   2243 GCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTC 2302
           | | ||| || |||| |          || |  |||||| ||| ||||     |  | ||
Db   2962 GAGTAGCCGCCGCCGCAGCTAGAGCCACCGCCGCCTCCGGAGTACTTGACGCCCCCACCG 2903

Qy   2303 TCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
            |||||  ||  |    || |   | ||| || ||
Db   2902 GAGCCGCCTCCACCGCCGCAGCTGGAACCACCTCC 2868


RESULT 23
PCT-US93-03993-1/c
; Sequence 1, Application PC/TUS9303993
;  GENERAL INFORMATION:
;    APPLICANT:  Roop, Dennis R.
;    APPLICANT:  Rothnagel, Joseph A.
;    APPLICANT:  Greenhalgh, David A.
;    TITLE OF INVENTION:  CONSTITUTIVE AND INDUCIBLE EPIDERMAL
;    TITLE OF INVENTION:  VECTOR SYSTEMS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/03993
;      FILING DATE:  19930428
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5405
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5325
;      TELEFAX:  713/651-5246
;      TELEX:  762829
```

```
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6530 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
PCT-US93-03993-1

  Query Match              1.1%;  Score 36.6;  DB 5;  Length 6530;
  Best Local Similarity   45.8%;  Pred. No. 5.5;
  Matches  126;  Conservative    0;  Mismatches  149;  Indels    0;  Gaps    0;

Qy   2063 CAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATC 2122
          | |  | || |  ||| ||| || |         ||||     |   || | || |
Db   3142 CCGGAGTACTTGACGCCCCCACCAGAGCCGCCACCGCCTCCACAGCTGGAGCCACCCCCG 3083

Qy   2123 TGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTG 2182
            | |          | || |   |  ||| |||   | |||| |  | ||    | |  |
Db   3082 GAGTAGCTGCCGCCGCCGCAGCTGGAGCCGCCGCCGCCTCCGGAGTAGCCACCTCCGCAG 3023

Qy   2183 AAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCC 2242
               || ||||||| || || ||        | | |          |  | |  | | |
Db   3022 CTGGAGCCGCCGCCGCCCCCAGAGTAGCCACCGCCGCAGCTAGAGCCACCACCGCCTCCG 2963

Qy   2243 GCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTC 2302
           | | ||| || |||| |          || |  |||||| ||| ||||     |  | ||
Db   2962 GAGTAGCCGCCGCCGCAGCTAGAGCCACCGCCGCCTCCGGAGTACTTGACGCCCCCACCG 2903

Qy   2303 TCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
            |||||  ||  |    || |   | ||| || ||
Db   2902 GAGCCGCCTCCACCGCCGCAGCTGGAACCACCTCC 2868


RESULT 24
US-08-895-590-43
; Sequence 43, Application US/08895590
; Patent No. 6207410
;  GENERAL INFORMATION:
;    APPLICANT:  Hall, Linda M.
;    APPLICANT:  Ren, Dejian
;    APPLICANT:  Zheng, Wei
;    APPLICANT:  Dubald, Manuel Marcel Paul
;    TITLE OF INVENTION:  Genes Encoding an Insect Calcium Channel
;    NUMBER OF SEQUENCES:  101
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS, LLP
;      STREET:  699 Prince Street
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22314-3187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/895,590
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/374,888
;      FILING DATE:  19-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McGowan, Malcolm M.
;      REGISTRATION NUMBER:  39,300
;      REFERENCE/DOCKET NUMBER:  022650-263
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
```

```
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1631 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  33..1631
US-08-895-590-43

  Query Match             1.1%;  Score 36.4;  DB 4;  Length 1631;
  Best Local Similarity   47.1%;  Pred. No. 3;
  Matches  112;  Conservative    0;  Mismatches  126;  Indels    0;  Gaps    0;

Qy    557 TCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAG 616
          ||||||| ||  |      |  ||   | ||||    ||    || |||||  |  |
Db    274 TCGACGACGATTCGCCGGTGCAGGGCGGAGCAGTTTCGTTCGAGATCCGGAATTCCTGC 333

Qy    617 AAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGAC 676
          | |||  | | |  | |       |        || ||| |||     |  ||| ||| |
Db    334 ACATCAAAGAGAAGTACGACAATAACTGGTGGATCGGCCGGCTGGTCAAGGAGGGCTGTG 393

Qy    677 CTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCA 736
            || |||    ||  |    |   || | |||| |||  ||     | |||  |    |||||
Db    394 ATGTGGGCTTCATACCCAGCCCCGCTAAGCTAGACAACATTCGCATGCAGAACCAAACCA 453

Qy    737 AGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 794
                |  || | ||| ||   | |    ||     |   |  |  || || || |    | ||
Db    454 GACCCTCAAGACTGTATGGCACAAAGGGCTCGTCGAGCGGGAATCTTGGTGCGGGCGG 511


RESULT 25
US-09-060-756-4/c
; Sequence 4, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
;  APPLICANT: Cole, Stewart
;  APPLICANT: Buchrieser-Brosch, Roland
;  APPLICANT: Gordon, Stephen
;  APPLICANT: Billault, Alain
;  TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
;  TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
;  TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
;  FILE REFERENCE: 3495-0169
;  CURRENT APPLICATION NUMBER: US/09/060,756
;  CURRENT FILING DATE: 1998-04-16
;  NUMBER OF SEQ ID NOS: 743
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 1280
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-060-756-4

  Query Match             1.1%;  Score 36.2;  DB 4;  Length 1280;
  Best Local Similarity   49.2%;  Pred. No. 3;
  Matches   95;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy   2179 GTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACT 2238
          | ||    | |  ||| ||| || | | | | |  | | |          || || |||
Db    882 GCTGCCGCCGACGCCACCGGCCCCGCCGGTGTCGCCGACGCTTAAGGCTGCGCCGCCTTC 823

Qy   2239 TTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAAC 2298
            |  ||  || || ||      |    |    ||  ||||  | ||  || | || |  |  |
Db    822 GACGCCGGCGCCGCCGCTACCCCATAACTGACCGCCGGTGCCGCCGTTCCCACCGGCGCC 763

Qy   2299 AGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGC 2358
                |   || ||||| | || ||| |  | || ||| ||   |||||     |||  |   |
Db    762 CCCGTTAGCGTTACCGCCGATTCCGCCGATGCCCCCGGCCCCGCCGAAGCCCAGCCAGCC 703

Qy   2359 CGGCTGAAAGCCG 2371
                  |  |      ||||
Db    702 GGCATTGCCGCCG 690


RESULT 26
US-08-471-033-24
; Sequence 24, Application US/08471033
; Patent No. 5770696
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  No. 5770696el Pesticidal Proteins and Strains
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-GEIGY Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,033
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  P-40,403
;      REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3/DIV7 - SQLv3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8582
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1399 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1386
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence for VIP2A(a) protein from AB78"
US-08-471-033-24

  Query Match             1.1%;  Score 36.2;  DB 1;  Length 1399;
  Best Local Similarity   45.1%;  Pred. No. 3.2;
  Matches  134;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy   2466 CAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGAC 2525
            || |||  ||||||    | |    ||     ||  |  |  |||        ||| || |  |
Db    339 CAGCATGGCCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398
```

```
Qy   2526 AGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAA 2585
           |    |||  ||| |||    |   || |||  |   |        | |    | |
Db    399 CGACAAGACCAACCTGAGCAACAGCATCATCACCTACAAGAACGTGGAGCCCACCACCAT 458

Qy   2586 AACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCT 2645
             | || ||  || |||   |  ||||         | |    | ||  || | | |  |
Db    459 CGGCTTCAACAAGAGCCTGACCGAGGGCAACACCATCAACAGCGACGCCATGGCCCAGTT 518

Qy   2646 CAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGA 2705
           ||| |  | ||||     || |   |  |||   |  || |  ||      ||||      |
Db    519 CAAGGAGCAGTTCCTGGACCGCGACATCAAGTTCGACAGCTACCTGGACACCCACCTGAC 578

Qy   2706 CGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
           ||  ||    ||    |||||    |  ||  | |  || || ||  ||  |    ||||
Db    579 CGCCCAGCAGGTGAGCAGCAAGGAGCGCGTGATCCTGAAGGTGACCGTCCCCAGCGG 635
```

RESULT 27
US-08-471-044-24
; Sequence 24, Application US/08471044
; Patent No. 5840868
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,044
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SQLv3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1386
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIP2A(a) protein from AB78"
US-08-471-044-24

```
  Query Match             1.1%;  Score 36.2;  DB 2;  Length 1399;
  Best Local Similarity   45.1%;  Pred. No. 3.2;
  Matches  134;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy   2466 CAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGAC 2525
           ||  |||  ||||||   | |    ||     ||  |  |  |||         ||| || | |
Db    339 CAGCATGGCCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398

Qy   2526 AGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAA 2585
           |    |||  ||| |||    |   || |||  |   |        | |    | |
Db    399 CGACAAGACCAACCTGAGCAACAGCATCATCACCTACAAGAACGTGGAGCCCACCACCAT 458

Qy   2586 AACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCT 2645
             | || ||  || |||   |  ||||         | |    | ||  || | | |  |
Db    459 CGGCTTCAACAAGAGCCTGACCGAGGGCAACACCATCAACAGCGACGCCATGGCCCAGTT 518

Qy   2646 CAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGA 2705
           ||| |  | ||||     || |   |  |||   |  || |  ||      ||||      |
Db    519 CAAGGAGCAGTTCCTGGACCGCGACATCAAGTTCGACAGCTACCTGGACACCCACCTGAC 578

Qy   2706 CGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
           ||  ||    ||    |||||    |  ||  | |  || || ||  ||  |    ||||
Db    579 CGCCCAGCAGGTGAGCAGCAAGGAGCGCGTGATCCTGAAGGTGACCGTCCCCAGCGG 635
```

RESULT 28
US-08-463-483A-24
; Sequence 24, Application US/08463483A
; Patent No. 5849870
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,483A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Spruill, W. Murray
;      REGISTRATION NUMBER:  32,943
;      REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8615
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1399 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1386
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence for VIP2A(a) protein from AB78"
US-08-463-483A-24

  Query Match             1.1%;  Score 36.2;  DB 2;  Length 1399;
  Best Local Similarity   45.1%;  Pred. No. 3.2;
  Matches  134;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy   2466 CAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGAC 2525
            || |||  ||||||   | |   ||     ||  | |  |||       ||| || | |
Db    339 CAGCATGGCCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398

Qy   2526 AGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAA 2585
           |       |||  ||| |||   |     || ||| |  |   |           | |    | |
Db    399 CGACAAGACCAACCTGAGCAACAGCATCATCACCTACAAGAACGTGGAGCCCACCACCAT 458

Qy   2586 AACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCT 2645
              |  || ||  || |||   |  ||||        | |      | ||  || | | |  |
Db    459 CGGCTTCAACAAGAGCCTGACCGAGGGCAACACCATCAACAGCGACGCCATGGCCCAGTT 518

Qy   2646 CAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGA 2705
            ||| |   | ||||      || |    |  |||    |  || |  ||      ||||      |
Db    519 CAAGGAGCAGTTCCTGGACCGCGACATCAAGTTCGACAGCTACCTGGACACCCACCTGAC 578

Qy   2706 CGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
            ||  ||    ||    |||||     |  ||   |  |  || || ||  ||   |    ||||
Db    579 CGCCCAGCAGGTGAGCAGCAAGGAGCGCGTGATCCTGAAGGTGACCGTCCCCAGCGG 635


RESULT 29
US-08-471-046A-24
; Sequence 24, Application US/08471046A
; Patent No. 5866326
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  Method For Isolating Vegetative Insecticidal
;    TITLE OF INVENTION:  Protein Genes
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5866326artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,046A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/463,483
;      FILING DATE:  05-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC1695/CIP3/DIV8 - SQLv4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1399 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1386
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence for VIP2A(a) protein from AB78"
US-08-471-046A-24

  Query Match             1.1%;  Score 36.2;  DB 2;  Length 1399;
  Best Local Similarity   45.1%;  Pred. No. 3.2;
  Matches  134;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy   2466 CAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGAC 2525
            || |||  ||||||   | |   ||     ||  | |  |||       ||| || | |
Db    339 CAGCATGGCCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398

Qy   2526 AGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAA 2585
           |       |||  ||| |||   |     || ||| |  |   |           | |    | |
Db    399 CGACAAGACCAACCTGAGCAACAGCATCATCACCTACAAGAACGTGGAGCCCACCACCAT 458

Qy   2586 AACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCT 2645
              |  || ||  || |||   |  ||||        | |      | ||  || | | |  |
Db    459 CGGCTTCAACAAGAGCCTGACCGAGGGCAACACCATCAACAGCGACGCCATGGCCCAGTT 518

Qy   2646 CAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGA 2705
            ||| |   | ||||      || |    |  |||    |  || |  ||      ||||      |
Db    519 CAAGGAGCAGTTCCTGGACCGCGACATCAAGTTCGACAGCTACCTGGACACCCACCTGAC 578

Qy   2706 CGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
            ||  ||    ||    |||||     |  ||   |  |  || || ||  ||   |    ||||
Db    579 CGCCCAGCAGGTGAGCAGCAAGGAGCGCGTGATCCTGAAGGTGACCGTCCCCAGCGG 635


RESULT 30
US-08-470-566B-24
; Sequence 24, Application US/08470566B
; Patent No. 5872212
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
```

```
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  No. 5872212el Pesticidal Proteins and Strains
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5872212artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,566B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/463,483
;      FILING DATE:  05-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC1695/CIP3/DIV4 - SQLv4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1399 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1386
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence for VIP2A(a) protein from AB78"
US-08-470-566B-24

  Query Match            1.1%;  Score 36.2;  DB 2;  Length 1399;
  Best Local Similarity  45.1%;  Pred. No. 3.2;
  Matches  134;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy   2466 CAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGAC 2525
           ||  |||  ||||||    | |    ||     ||  | |  |||        ||| || | |
Db    339 CAGCATGGCCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398

Qy   2526 AGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAA 2585
             |      |||  ||| |||    |    || ||| |   |    |          | |    | |
Db    399 CGACAAGACCAACCTGAGCAACAGCATCATCACCTACAAGAACGTGGAGCCCACCACCAT 458

Qy   2586 AACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCT 2645
               | || ||  || |||   |  ||||        | |     | ||  || | | |   |
Db    459 CGGCTTCAACAAGAGCCTGACCGAGGGCAACACCATCAACAGCGACGCCATGGCCCAGTT 518

Qy   2646 CAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGA 2705
           ||| |   | ||||     || |    |   |||    |   || |  ||      ||||      |
Db    519 CAAGGAGCAGTTCCTGGACCGCGACATCAAGTTCGACAGCTACCTGGACACCCACCTGAC 578

Qy   2706 CGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
           ||  ||    ||    |||||     |  ||   | |   || || ||   ||   |    ||||
Db    579 CGCCCAGCAGGTGAGCAGCAAGGAGCGCGTGATCCTGAAGGTGACCGTCCCCAGCGG 635


RESULT 31
US-08-469-334-24
; Sequence 24, Application US/08469334
; Patent No. 5990383
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  No. 5990383el Pesticidal Proteins and Strains
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-GEIGY Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,334
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/463,483
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Spruill, W. Murray
;      REGISTRATION NUMBER:  32,943
;      REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8615
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1399 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1386
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence for VIP2A(a) protein from AB78"
```

```
US-08-469-334-24

  Query Match             1.1%;  Score 36.2;  DB 2;  Length 1399;
  Best Local Similarity   45.1%;  Pred. No. 3.2;
  Matches  134;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy     2466 CAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGAC 2525
              ||  |||  ||||||    | |    ||      ||  | |  |||         ||| || | |
Db      339 CAGCATGGCCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398

Qy     2526 AGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAA 2585
                |       |||  ||| |||    |    || ||| |   |     |          | |   | |
Db      399 CGACAAGACCAACCTGAGCAACAGCATCATCACCTACAAGAACGTGGAGCCCACCACCAT 458

Qy     2586 AACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCT 2645
                 |  || ||  || |||   |  ||||          | |     | ||  || | | | |
Db      459 CGGCTTCAACAAGAGCCTGACCGAGGGCAACACCATCAACAGCGACGCCATGGCCCAGTT 518

Qy     2646 CAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGA 2705
              ||| |   | ||||     || |    |  |||    |  || |  ||      ||||        |
Db      519 CAAGGAGCAGTTCCTGGACCGCGACATCAAGTTCGACAGCTACCTGGACACCCACCTGAC 578

Qy     2706 CGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
              ||  ||    ||    |||||      |  ||  | |  || || ||  ||  |     ||||
Db      579 CGCCCAGCAGGTGAGCAGCAAGGAGCGCGTGATCCTGAAGGTGACCGTCCCCAGCGG 635


RESULT 32
US-09-300-529-24
; Sequence 24, Application US/09300529
; Patent No. 6066783
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  Genes Encoding Insecticidal Proteins
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6066783artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/300,529
;      FILING DATE:  TBA
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/469,334
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/463,483
;      FILING DATE:  05-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  S-19506L
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1399 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1386
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence for VIP2A(a) protein from AB78"
US-09-300-529-24

  Query Match             1.1%;  Score 36.2;  DB 3;  Length 1399;
  Best Local Similarity   45.1%;  Pred. No. 3.2;
  Matches  134;  Conservative    0;  Mismatches  163;  Indels    0;  Gaps    0;

Qy     2466 CAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGAC 2525
              ||  |||  ||||||    | |    ||      ||  | |  |||         ||| || | |
Db      339 CAGCATGGCCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTT 398

Qy     2526 AGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAA 2585
                |       |||  ||| |||    |    || ||| |   |     |          | |   | |
Db      399 CGACAAGACCAACCTGAGCAACAGCATCATCACCTACAAGAACGTGGAGCCCACCACCAT 458

Qy     2586 AACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCT 2645
                 |  || ||  || |||   |  ||||          | |     | ||  || | | | |
Db      459 CGGCTTCAACAAGAGCCTGACCGAGGGCAACACCATCAACAGCGACGCCATGGCCCAGTT 518

Qy     2646 CAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGA 2705
              ||| |   | ||||     || |    |  |||    |  || |  ||      ||||        |
Db      519 CAAGGAGCAGTTCCTGGACCGCGACATCAAGTTCGACAGCTACCTGGACACCCACCTGAC 578

Qy     2706 CGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
              ||  ||    ||    |||||      |  ||  | |  || || ||  ||  |     ||||
Db      579 CGCCCAGCAGGTGAGCAGCAAGGAGCGCGTGATCCTGAAGGTGACCGTCCCCAGCGG 635


RESULT 33
US-08-162-475A-3/c
; Sequence 3, Application US/08162475A
; Patent No. 5656474
;  GENERAL INFORMATION:
;    APPLICANT:  Zohreh Tabaeizadeh
;    TITLE OF INVENTION:  A novel endochitinase gene
;    TITLE OF INVENTION:  induced by osmotic stress and abscisic acid isolated
;    TITLE OF INVENTION:  from the wild tomato Lycopersicon chilense Dun.
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;      STREET:  1100 New York Avenue, N.W., 9th Floor
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 720 kb diskette
;      COMPUTER:  IBM PS/2, Model 30
;      OPERATING SYSTEM:  PC-DOS 3.30
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/162,475A
;      FILING DATE:  December 7, 1993
;      CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, Paul E. Jr.
;      REGISTRATION NUMBER:  32011
;      REFERENCE/DOCKET NUMBER:  PEW/3122/204351
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  935 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:  Lycopersicon chilense
US-08-162-475A-3

  Query Match             1.1%;  Score 36;  DB 1;  Length 935;
  Best Local Similarity   55.6%;  Pred. No. 2.9;
  Matches   69;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy   1443 TACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGG 1502
          || |||| | ||||  | ||  |  |||| |  || |  ||||     || | |  |
Db    316 TAACAAAGCAATATCCTCCAGTGAATGTTCCTGCACTACCACCATTAGTTTCATGAGATG 257

Qy   1503 CCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAA 1562
            ||| | |||||||| |||      || |     |  | | | || ||||  ||    |||
Db    256 TTTGACCGAAAAAGGCAGCAATTTCCTTCTTACGTGCAGTATCATCACCAGTAGTACCAA 197

Qy   1563 AACC 1566
          ||||
Db    196 AACC 193


RESULT 34
US-08-459-448A-26
; Sequence 26, Application US/08459448A
; Patent No. 5859336
;  GENERAL INFORMATION:
;    APPLICANT:  Koziel, Michael G.
;    APPLICANT:  Desai, Nalini M.
;    APPLICANT:  Lewis, Kelly S.
;    APPLICANT:  Kramer, Vance C.
;    APPLICANT:  Warren, Gregory W.
;    APPLICANT:  Evola, Stephen V.
;    APPLICANT:  Crossland, Lyle D.
;    APPLICANT:  Wright, Martha S.
;    APPLICANT:  Merlin, Ellis J.
;    APPLICANT:  Launis, Karen L.
;    APPLICANT:  Rothstein, Steven J.
;    APPLICANT:  Bowman, Cindy G.
;    APPLICANT:  Dawson, John L.
;    APPLICANT:  Dunder, Erik M.
;    APPLICANT:  Pace, Gary M.
;    APPLICANT:  Suttie, Janet L.
;    TITLE OF INVENTION:  SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY IN MAIZE
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5859336artis Corporation
;      STREET:  Patent & Trademark Dept., 520 White Plains
;      STREET:  Rd., POB 2005
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,448A
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/951,715
;      FILING DATE:  25-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/772,027
;      FILING DATE:  04-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  40403
;      REFERENCE/DOCKET NUMBER:  CGC 1577/CIP/DIV4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919)541-8582
;      TELEFAX:  (919)541-8689
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4162 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1418..1427
;      OTHER INFORMATION:  /note= "start of mRNA"
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1481..2366
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2367..2451
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2452..2602
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2603..2690
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2691..2804
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2805..2906
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2907..3075
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3076..3177
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3178..3304
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3305..3398
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3399..3498
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3499..3713
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3714..3811
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1..1477
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  /partial
;      OTHER INFORMATION:  /function= "pollen-specific promoter region"
;      OTHER INFORMATION:  /evidence= EXPERIMENTAL
```

```
US-08-459-448A-26

  Query Match             1.1%;  Score 35.6;  DB 2;  Length 4162;
  Best Local Similarity   54.6%;  Pred. No. 8.3;
  Matches   71;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy       38 AACCGGCTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCG 97
             ||| |||  |     |||  || || ||  |  |   | ||||   |   || || ||||
Db     1581 AACGGGCGTCGTCCTCGTCCGCGGTGGCTGCTGCCGCTGCTGCTGCCGGTGGTGGTGGCG 1640

Qy       98 GCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAG 157
             |||||||  |   ||| | ||       | |||||| || |   | ||  |  ||||| |
Db     1641 GCGGCACGACGAAGCCGGCCTCACCCACCGGCGGCGCCAGGGCCAGCTCCGGCAGCAAAC 1700

Qy      158 CAACAACAGC 167
             |  |  | ||
Db     1701 CGGCGGCGGC 1710


RESULT 35
US-08-459-595A-26
; Sequence 26, Application US/08459595A
; Patent No. 6018104
;  GENERAL INFORMATION:
;    APPLICANT:  Koziel, Michael G.
;    APPLICANT:  Desai, Nalini M.
;    APPLICANT:  Lewis, Kelly S.
;    APPLICANT:  Kramer, Vance C.
;    APPLICANT:  Warren, Gregory W.
;    APPLICANT:  Evola, Stephen V.
;    APPLICANT:  Crossland, Lyle D.
;    APPLICANT:  Wright, Martha S.
;    APPLICANT:  Merlin, Ellis J.
;    APPLICANT:  Launis, Karen L.
;    APPLICANT:  Rothstein, Steven J.
;    APPLICANT:  Bowman, Cindy G.
;    APPLICANT:  Dawson, John L.
;    APPLICANT:  Dunder, Erik M.
;    APPLICANT:  Pace, Gary M.
;    APPLICANT:  Suttie, Janet L.
;    TITLE OF INVENTION:  SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY IN MAIZE
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018104artis Corporation
;      STREET:  Patent & Trademark Dept., 520 White Plains
;      STREET:  Rd., POB 2005
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,595A
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/951,715
;      FILING DATE:  25-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/772,027
;      FILING DATE:  04-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  40403
;      REFERENCE/DOCKET NUMBER:  CGC 1577/CIP/DIV3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919)541-8582
;      TELEFAX:  (919)541-8689
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4162 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1418..1427
;      OTHER INFORMATION:  /note= "start of mRNA"
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1481..2366
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2367..2451
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2452..2602
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2603..2690
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2691..2804
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2805..2906
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2907..3075
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3076..3177
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3178..3304
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3305..3398
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3399..3498
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3499..3713
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3714..3811
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1..1477
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  /partial
;      OTHER INFORMATION:  /function= "pollen-specific promoter region"
;      OTHER INFORMATION:  /evidence= EXPERIMENTAL
US-08-459-595A-26

  Query Match             1.1%;  Score 35.6;  DB 3;  Length 4162;
  Best Local Similarity   54.6%;  Pred. No. 8.3;
  Matches   71;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy       38 AACCGGCTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCG 97
             ||| |||  |     |||  || || ||  |  |   | ||||   |   || || ||||
Db     1581 AACGGGCGTCGTCCTCGTCCGCGGTGGCTGCTGCCGCTGCTGCTGCCGGTGGTGGTGGCG 1640

Qy       98 GCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAG 157
             |||||||  |   ||| | ||       | |||||| || |   | ||  |  ||||| |
Db     1641 GCGGCACGACGAAGCCGGCCTCACCCACCGGCGGCGCCAGGGCCAGCTCCGGCAGCAAAC 1700

Qy      158 CAACAACAGC 167
```

```
             |  |  | ||
Db   1701 CGGCGGCGGC 1710


RESULT 36
US-08-459-504B-26
; Sequence 26, Application US/08459504B
; Patent No. 6075185
;  GENERAL INFORMATION:
;    APPLICANT:  Koziel, Michael G.
;    APPLICANT:  Desai, Nalini M.
;    APPLICANT:  Lewis, Kelly S.
;    APPLICANT:  Kramer, Vance C.
;    APPLICANT:  Warren, Gregory W.
;    APPLICANT:  Evola, Stephen V.
;    APPLICANT:  Crossland, Lyle D.
;    APPLICANT:  Wright, Martha S.
;    APPLICANT:  Merlin, Ellis J.
;    APPLICANT:  Launis, Karen L.
;    APPLICANT:  Rothstein, Steven J.
;    APPLICANT:  Bowman, Cindy G.
;    APPLICANT:  Dawson, John L.
;    APPLICANT:  Dunder, Erik M.
;    APPLICANT:  Pace, Gary M.
;    APPLICANT:  Suttie, Janet L.
;    TITLE OF INVENTION:  SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY IN MAIZE
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6075185artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,504B
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/459,595
;      FILING DATE:  02-JUN-1995
;      APPLICATION NUMBER:  US 07/951,715
;      FILING DATE:  25-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/772,027
;      FILING DATE:  04-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC1577/CIP/DIV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919)541-8587
;      TELEFAX:  (919)541-8689
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4162 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1418..1427
;      OTHER INFORMATION:  /note= "start of mRNA"
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1481..2366
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2367..2451
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2452..2602
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2603..2690
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2691..2804
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2805..2906
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2907..3075
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3076..3177
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3178..3304
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3305..3398
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3399..3498
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3499..3713
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3714..3811
US-08-459-504B-26

  Query Match             1.1%;  Score 35.6;  DB 3;  Length 4162;
  Best Local Similarity   54.6%;  Pred. No. 8.3;
  Matches   71;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy     38 AACCGGCTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCG 97
          ||| |||  |      |||  || || ||  |  |   | ||||   |    || || ||||
Db   1581 AACGGGCGTCGTCCTCGTCCGCGGTGGCTGCTGCCGCTGCTGCTGCCGGTGGTGGTGGCG 1640

Qy     98 GCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAG 157
          |||||||  |   ||| | ||      | |||||| || |    | ||   |  ||||| |
Db   1641 GCGGCACGACGAAGCCGGCCTCACCCACCGGCGGCGCCAGGGCCAGCTCCGGCAGCAAAC 1700

Qy    158 CAACAACAGC 167
          |  |  | ||
Db   1701 CGGCGGCGGC 1710


RESULT 37
US-08-459-444-26
; Sequence 26, Application US/08459444A
; Patent No. 6121014
;  GENERAL INFORMATION:
;       APPLICANT: Koziel, Michael G.
;                  Desai, Nalini M.
;                  Lewis, Kelly S.
;                  Kramer, Vance C.
;                  Warren, Gregory W.
;                  Evola, Stephen V.
;                  Crossland, Lyle D.
;                  Wright, Martha S.
;                  Merlin, Ellis J.
;                  Launis, Karen L.
;       TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
;                           NUCLEIC ACID CODING SEQUENCE
```

```
;       NUMBER OF SEQUENCES: 94
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
;            STREET: 3054 Cornwallis Road
;            CITY: Research Triangle Park
;            STATE: NC
;            COUNTRY: USA
;            ZIP: 27709
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/459,444A
;            FILING DATE: 02-Jun-1995
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 07/951,715
;            FILING DATE: 25-SEP-1992
;            APPLICATION NUMBER: US 07/772,027
;            FILING DATE: 04-OCT-1991
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Meigs, J. Timothy
;            REGISTRATION NUMBER: 38,241
;            REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (919)541-8587
;            TELEFAX: (919)541-8689
;  INFORMATION FOR SEQ ID NO: 26:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 4162 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       HYPOTHETICAL: NO
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION:  1418..1427
;            OTHER INFORMATION:  /note= "start of mRNA"
;       FEATURE:
;            NAME/KEY:  exon
;            LOCATION:  1481..2366
;       FEATURE:
;            NAME/KEY:  intron
;            LOCATION:  2367..2451
;       FEATURE:
;            NAME/KEY:  exon
;            LOCATION:  2452..2602
;       FEATURE:
;            NAME/KEY:  intron
;            LOCATION:  2603..2690
;       FEATURE:
;            NAME/KEY:  exon
;            LOCATION:  2691..2804
;       FEATURE:
;            NAME/KEY:  intron
;            LOCATION:  2805..2906
;       FEATURE:
;            NAME/KEY:  exon
;            LOCATION:  2907..3075
;       FEATURE:
;            NAME/KEY:  intron
;            LOCATION:  3076..3177
;       FEATURE:
;            NAME/KEY:  exon
;            LOCATION:  3178..3304
;       FEATURE:
;            NAME/KEY:  intron
;            LOCATION:  3305..3398
;       FEATURE:
;            NAME/KEY:  exon
;            LOCATION:  3399..3498
;       FEATURE:
;            NAME/KEY:  intron
;            LOCATION:  3499..3713
;       FEATURE:
;            NAME/KEY:  exon
;            LOCATION:  3714..3811
;       SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-459-444-26

  Query Match             1.1%;  Score 35.6;  DB 3;  Length 4162;
  Best Local Similarity   54.6%;  Pred. No. 8.3;
  Matches   71;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy         38 AACCGGCTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCG 97
              ||| |||  |     |||  || || ||  |  |   | ||||   |   || || ||||
Db       1581 AACGGGCGTCGTCCTCGTCCGCGGTGGCTGCTGCCGCTGCTGCTGCCGGTGGTGGTGGCG 1640

Qy         98 GCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAG 157
              |||||||  |    ||| | ||        | |||||| || |    | ||  |  ||||| |
Db       1641 GCGGCACGACGAAGCCGGCCTCACCCACCGGCGGCGCCAGGGCCAGCTCCGGCAGCAAAC 1700

Qy        158 CAACAACAGC 167
              |  |  | ||
Db       1701 CGGCGGCGGC 1710


RESULT 38
US-09-547-422-26
; Sequence 26, Application US/09547422
; Patent No. 6320100
;  GENERAL INFORMATION:
;       APPLICANT: Koziel, Michael G.
;                  Desai, Nalini M.
;                  Lewis, Kelly S.
;                  Kramer, Vance C.
;                  Warren, Gregory W.
;                  Evola, Stephen V.
;                  Crossland, Lyle D.
;                  Wright, Martha S.
;                  Merlin, Ellis J.
;                  Launis, Karen L.
;       TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;                           INSECTICIDAL ACTIVITY IN MAIZE
;       NUMBER OF SEQUENCES: 94
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.
;            STREET: 3054 Cornwallis Road
;            CITY: Research Triangle Park
;            STATE: NC
;            COUNTRY: USA
;            ZIP: 27709
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/547,422
;            FILING DATE: 11-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/459,595
;            FILING DATE: 02-JUN-1995
;            APPLICATION NUMBER: US 07/951,715
;            FILING DATE: 25-SEP-1992
;            APPLICATION NUMBER: US 07/772,027
;            FILING DATE: 04-OCT-1991
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Meigs, J. Timothy
;            REGISTRATION NUMBER: 38,241
;            REFERENCE/DOCKET NUMBER: S-18805H
;       TELECOMMUNICATION INFORMATION:
```

```
;             TELEPHONE: (919)541-8587
;             TELEFAX: (919)541-8689
;   INFORMATION FOR SEQ ID NO: 26:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 4162 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        HYPOTHETICAL: NO
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION:  1418..1427
;             OTHER INFORMATION:  /note= "start of mRNA"
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  1481..2366
;        FEATURE:
;             NAME/KEY:  intron
;             LOCATION:  2367..2451
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  2452..2602
;        FEATURE:
;             NAME/KEY:  intron
;             LOCATION:  2603..2690
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  2691..2804
;        FEATURE:
;             NAME/KEY:  intron
;             LOCATION:  2805..2906
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  2907..3075
;        FEATURE:
;             NAME/KEY:  intron
;             LOCATION:  3076..3177
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  3178..3304
;        FEATURE:
;             NAME/KEY:  intron
;             LOCATION:  3305..3398
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  3399..3498
;        FEATURE:
;             NAME/KEY:  intron
;             LOCATION:  3499..3713
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  3714..3811
;        SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-547-422-26

  Query Match             1.1%;  Score 35.6;  DB 4;  Length 4162;
  Best Local Similarity   54.6%;  Pred. No. 8.3;
  Matches   71;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy       38 AACCGGCTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCG 97
            ||| |||  |      |||  || || ||  |  |    | ||||   |    || || ||||
Db     1581 AACGGGCGTCGTCCTCGTCCGCGGTGGCTGCTGCCGCTGCTGCTGCCGGTGGTGGCG 1640

Qy       98 GCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAG 157
            |||||||  |    ||| | ||       | |||||| || |    | ||   |   ||||| |
Db     1641 GCGGCACGACGAAGCCGGCCTCACCCACCGGCGGCGCCAGGGCCAGCTCCGGCAGCAAAC 1700

Qy      158 CAACAACAGC 167
            |  |   | ||
Db     1701 CGGCGGCGGC 1710
```

```
RESULT 39
US-07-951-715A-26
; Sequence 26, Application US/07951715A
; Patent No. 5625136
;  GENERAL INFORMATION:
;    APPLICANT:  Koziel, Michael G.
;    APPLICANT:  Desai, Nalini M.
;    APPLICANT:  Lewis, Kelly S.
;    APPLICANT:  Kramer, Vance C.
;    APPLICANT:  Warren, Gregory W.
;    APPLICANT:  Evola, Stephen V.
;    APPLICANT:  Crossland, Lyle D.
;    APPLICANT:  Wright, Martha S.
;    APPLICANT:  Merlin, Ellis J.
;    APPLICANT:  Launis, Karen L.
;    APPLICANT:  Rothstein, Steven J.
;    APPLICANT:  Bowman, Cindy G.
;    APPLICANT:  Dawson, John L.
;    APPLICANT:  Dunder, Erik M.
;    APPLICANT:  Pace, Gary M.
;    APPLICANT:  Suttie, Janet L.
;    TITLE OF INVENTION:  SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY IN MAIZE
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-GEIGY Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/951,715A
;      FILING DATE:  25-SEP-1992
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/772,027
;      FILING DATE:  04-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Spruill, W. Murray
;      REGISTRATION NUMBER:  32,943
;      REFERENCE/DOCKET NUMBER:  S-18805/A/CGC 1577/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919)541-8615
;      TELEFAX:  (919)541-8689
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4165 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1416..1425
;      OTHER INFORMATION:  /note= "start of mRNA"
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1481..2366
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2367..2449
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2450..2602
;    FEATURE:
;      NAME/KEY:  intron
```

```
;      LOCATION:  2603..2688
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2689..2804
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2805..2906
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2907..3074
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3075..3177
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3178..3305
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3306..3397
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3398..3497
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  3498..3712
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3713..3811
US-07-951-715A-26

  Query Match             1.1%;  Score 35.6;  DB 1;  Length 4165;
  Best Local Similarity   54.6%;  Pred. No. 8.3;
  Matches   71;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy     38 AACCGGCTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCG 97
          ||| |||  |     |||  || || ||  |  |    | ||||    |    || || ||||
Db   1579 AACGGGCGTCGTCCTCGTCCGCGGTGGCTGCTGCCGCTGCTGCTGCCGGTGGTGGTGGCG 1638

Qy     98 GCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAG 157
          |||||||  |     ||| |  ||       | |||||| || |     | ||   |  ||||| |
Db   1639 GCGGCACGACGAAGCCGGCCTCACCCACCGGCGGCGCCAGGGCCAGCTCCGGCAGCAAAC 1698

Qy    158 CAACAACAGC 167
          |  |   | ||
Db   1699 CGGCGGCGGC 1708


RESULT 40
US-09-105-537-34
; Sequence 34, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT: Liu, H.
;  APPLICANT: Xue, Y.
;  APPLICANT: Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/105,537A
;  CURRENT FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 34
;   LENGTH: 4689
;   TYPE: DNA
;   ORGANISM: Streptomyces venezuelae
US-09-105-537-34

  Query Match             1.1%;  Score 35.6;  DB 4;  Length 4689;
  Best Local Similarity   45.4%;  Pred. No. 8.8;
  Matches  128;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy   2077 GCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATG 2136
          || | |  |||     |||||  | |         || |||||        ||  ||
Db   3406 GCCCACCACGGAGCCGAACACCTCCTCCTCGTCAGCCGCAGCGGCGAACAAGCCCCCGGA 3465

Qy   2137 GTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCC 2196
          | |   |       || |||||     ||||| |   ||      ||     |  || |||  |
Db   3466 GCCACCCAACTCACCGCCGAACTCACCGCATCGGGCGCCCGCGTCACCATCGCCGCCTGC 3525

Qy   2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
          |||  |  | ||       |   | ||| ||| ||  |   | |    |   ||| ||| |  |  |   ||
Db   3526 GACGTCGCCGACCCCCACGCCATGCGCACCCTCCTCGACGCCATCCCCGCCGAGACGCCC 3585

Qy   2257 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTC 2316
           |       |         | ||| ||| || |      |  | |     || | ||  |||| |
Db   3586 CTCACCGCCGTCGTCCACACCGCCGGCGCACCGGGCGGCGATCCGCTGGACGTCACCGGC 3645

Qy   2317 TATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGC 2358
             |  ||||    ||   |   ||   |·||   |  |   |  |   ||
Db   3646 CCGGAGGACATCGCCCGCATCCTGGGCGCGAAGACGAGCGGC 3687

Search completed: January 27, 2003, 21:43:24
Job time : 9300 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 16:39:33 ; Search time 434 Seconds
(without alignments)
16625.352 Million cell updates/sec

Title: US-09-830-433A-7
Perfect score: 3204
Sequence: 1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3204	100.0	3204	21	AAA15299	DNA encoding a pol
c	2	3062.2	95.6	13423	21	AAA81527	N. meningitidis pa
	3	3062.2	95.6	349980	21	AAF21612	Neisseria meningit
	4	2986.2	93.2	4218	22	AAS43905	Neisseria meningit
	5	2986.2	93.2	4218	22	AAD17058	N. meningitidis st
	6	2983.8	93.1	4335	22	AAS43899	Neisseria meningit
	7	2983.8	93.1	4335	22	AAD17052	N. meningitidis st
	8	2983.2	93.1	3939	22	AAS43880	Neisseria meningit
	9	2983.2	93.1	3939	22	AAD17045	N. meningitidis st
	10	2983.2	93.1	4170	22	AAS43902	Neisseria meningit
	11	2983.2	93.1	4170	22	AAD17055	N. meningitidis st
	12	2971.8	92.8	3939	22	AAS43875	Neisseria meningit
	13	2971.8	92.8	3939	22	AAD17040	N. meningitidis st
	14	2971.8	92.8	4179	22	AAS43877	Neisseria meningit
	15	2971.8	92.8	4179	22	AAD17042	N. meningitidis st
	16	2971.8	92.8	4344	22	AAS43876	Neisseria meningit
	17	2971.8	92.8	4425	22	AAS43874	Neisseria meningit
	18	2971.8	92.8	4425	22	AAD17039	N. meningitidis st
	19	1946.6	60.8	3798	22	AAD17041	N. meningitidis st
	20	1349	42.1	1365	21	AAZ53301	Neisseria meningit
	21	1342.6	41.9	1365	21	AAZ53300	Neisseria meningit
	22	1288.2	40.2	1365	21	AAZ53299	Neisseria gonorrhe
	23	494.2	15.4	526	21	AAZ53298	Neisseria meningit
	24	470.6	14.7	528	21	AAZ53297	Neisseria meningit
	25	433.4	13.5	537	21	AAZ53296	Neisseria gonorrhe
	26	270.8	8.5	286	19	AAV03603	Neisseria meningit
	27	270.8	8.5	286	21	AAA15396	Genomic fragment o
	28	74.4	2.3	577	21	AAA82003	N. meningitidis pa
c	29	65.4	2.0	476	21	AAA81823	N. meningitidis pa
	30	62	1.9	10732	21	AAA10594	Gene encoding a su
	31	49.6	1.5	913	18	AAT91476	Mycobacterium tube
	32	49.6	1.5	913	18	AAT91413	Mycobacterium tube
	33	49.6	1.5	913	19	AAV64462	M. tuberculosis im
	34	49.6	1.5	913	19	AAV44354	Mycobacterium tube
	35	49.6	1.5	913	20	AAZ19264	M. tuberculosis an
	36	49.6	1.5	913	20	AAZ19052	M. tuberculosis re
	37	49.6	1.5	4403765	22	AAI99683	Mycobacterium tube
	38	49.6	1.5	4411529	22	AAI99682	Mycobacterium tube
	39	48.4	1.5	750	24	ABK74700	Bacillus lichenifo
c	40	45.2	1.4	4403765	22	AAI99683	Mycobacterium tube
c	41	45.2	1.4	4411529	22	AAI99682	Mycobacterium tube
	42	45	1.4	4384	22	AAF57288	M. bovis Dalton 2d
	43	41.8	1.3	3946	18	AAT93610	Mycobacterium tube
	44	41.4	1.3	2949	21	AAD00024	Degenerate DNA enc
	45	41.4	1.3	2949	24	AAD37455	Human Zfsta2 degen

ALIGNMENTS

RESULT 1
AAA15299
ID AAA15299 standard; DNA; 3204 BP.
XX
AC AAA15299;
XX
DT 04-SEP-2000 (first entry)
XX
DE DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX
KW Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1..3204
FT /*tag= a
XX
PN WO200026375-A2.
XX
PD 11-MAY-2000.
XX
PF 28-OCT-1999; 99WO-FR02643.
XX
PR 30-OCT-1998; 98FR-0013693.
XX
PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
PI Tinsley C, Perrin A;
XX

Nov 2000

```
DR   WPI; 2000-365622/31.
DR   P-PSDB; AAY93268.
XX
PT   New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT   or preventative vaccines and for diagnosis  -
XX
PS   Claim 2; Page 61-65; 187pp; French.
XX
CC   The present sequence encodes a protein that is specific for pathogenic
CC   strains of Neisseria. The polynucleotides, polypeptides, or their
CC   antigenic fragments, are used in vaccines to treat or protect against
CC   Neisseria infections, particularly by N. meningitidis. The
CC   polynucleotide sequence is also used for recombinant production of
CC   the polypeptide and to produce attenuated Neisseria strains that
CC   overexpress it, or express it in a non-toxic mutant form.
XX
SQ   Sequence 3204 BP; 831 A; 909 C; 899 G; 565 T; 0 other;

  Query Match             100.0%;  Score 3204;  DB 21;  Length 3204;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy       61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120

Qy      121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180

Qy      181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240

Qy      241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300

Qy      301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360

Qy      361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420

Qy      421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480

Qy      481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540

Qy      541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600

Qy      601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660

Qy      661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720

Qy      721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780

Qy      781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840

Qy      841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900

Qy      901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960

Qy      961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020

Qy     1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080

Qy     1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140

Qy     1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200

Qy     1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260

Qy     1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320

Qy     1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380

Qy     1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440

Qy     1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500

Qy     1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560

Qy     1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620

Qy     1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680

Qy     1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740

Qy     1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800

Qy     1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
```

```
Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180

Qy   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||
Db   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
```

```
RESULT 2
AAA81527/c
ID   AAA81527 standard; DNA; 13423 BP.
XX
AC   AAA81527;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_74 SEQ ID NO:74.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1439-1443; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
```

```
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 13423 BP; 2916 A; 3336 C; 3842 G; 3329 T; 0 other;

  Query Match             95.6%;  Score 3062.2;  DB 21;  Length 13423;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3156;  Conservative    0;  Mismatches   48;  Indels   45;  Gaps    2;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
Db      10889 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGACTGCCATGGCGTTAGCT 10830

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
Db      10829 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 10770

Qy        118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
Db      10769 TTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 10710

Qy        178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
Db      10709 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 10650

Qy        238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 297
Db      10649 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 10590

Qy        298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 357
Db      10589 CTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 10530

Qy        358 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGC 417
Db      10529 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC 10470

Qy        418 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 477
Db      10469 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 10410

Qy        478 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 537
Db      10409 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 10350

Qy        538 AAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 597
Db      10349 AAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 10290

Qy        598 GATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGG 657
Db      10289 GATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGG 10230

Qy        658 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 717
Db      10229 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 10170

Qy        718 AATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTC 777
Db      10169 AATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTC 10110

Qy        778 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGAGGGCA 837
Db      10109 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGAGGGCA 10050

Qy        838 GGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 897
Db      10049 GGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 9990

Qy        898 GCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 957
Db       9989 GACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 9930

Qy        958 GGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAAT 1017
Db       9929 GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAAT 9870

Qy       1018 GACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAA 1077
Db       9869 GACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAA 9810

Qy       1078 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA----------- 1126
Db       9809 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATG 9750

Qy       1127 ------------------------------ATGGCTCCAACCATTGCGGAATTACTGCC 1155
Db       9749 TATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCC 9690

Qy       1156 ATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 1215
Db       9689 ATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 9630

Qy       1216 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 1275
Db       9629 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 9570

Qy       1276 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAG 1335
Db       9569 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAG 9510

Qy       1336 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 1395
Db       9509 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 9450

Qy       1396 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 1455
Db       9449 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 9390

Qy       1456 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 1515
Db       9389 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 9330

Qy       1516 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 1575
Db       9329 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 9270

Qy       1576 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 1635
Db       9269 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 9210

Qy       1636 GCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTAT 1695
Db       9209 GCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT 9150

Qy       1696 CTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAG 1755
Db       9149 CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAG 9090

Qy       1756 CTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 1815
Db       9089 CTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 9030

Qy       1816 GCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAAC 1875
Db       9029 GCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAAC 8970
```

```
Qy   1876 CGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTC 1935
           ||||||||| ||||||||||||||||||||||||||||||||||| |||||||||||||
Db   8969 AGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTC 8910

Qy   1936 TTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 1995
          ||||||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db   8909 TTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 8850

Qy   1996 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 2055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8849 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 8790

Qy   2056 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 2115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8789 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 8730

Qy   2116 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 2175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8729 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 8670

Qy   2176 ACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 2235
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   8669 ACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 8610

Qy   2236 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 2295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8609 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 8550

Qy   2296 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 2355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8549 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 8490

Qy   2356 CGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATC 2415
          ||||| ||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db   8489 CGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATC 8430

Qy   2416 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 2475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8429 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 8370

Qy   2476 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 2535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8369 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 8310

Qy   2536 ACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 2595
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   8309 ACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 8250

Qy   2596 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 2655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8249 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 8190

Qy   2656 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 2715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8189 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 8130

Qy   2716 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 2775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8129 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 8070

Qy   2776 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 2835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8069 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 8010

Qy   2836 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 2895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8009 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 7950

Qy   2896 ACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 2955
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7949 ACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 7890

Qy   2956 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 3015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7889 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 7830

Qy   3016 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGC 3075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7829 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGT 7770

Qy   3076 CTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 3135
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   7769 CTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 7710

Qy   3136 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 3195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7709 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 7650

Qy   3196 CGGTTCTGA 3204
          |||||||||
Db   7649 CGGTTCTGA 7641
```

```
RESULT 3
AAF21612
ID   AAF21612 standard; DNA; 349980 BP.
XX
AC   AAF21612;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,   Hickey E,   Peterson J,   Tettelin H,   Venter JC,   Masignani V;
PI   Galeotti C,   Mora M,   Ratti G,   Scarselli M,   Scarlato V,   Rappuoli R;
PI   Frazer CM,   Grandi G;
XX
DR   WPI; 2000-647603/62.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own it was split into 8
CC   sequences which overlap each other at the beginning and end of each
CC   sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC   the beginning  of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC   the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC   Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC   AAF21606 represent PCR primers which are used in the exemplification of
CC   the present invention. The NMB genome and fragments from it have
CC   antibacterial activity, and can be used in vaccines and gene therapy.
CC   Neisseria nucleic acids, proteins and/or antibodies which binds to the
```

```
CC   proteins can be used in compositions for treating or preventing infection
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

  Query Match             95.6%;  Score 3062.2;  DB 21;  Length 349980;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3156;  Conservative    0;  Mismatches   48;  Indels   45;  Gaps    2;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
Db     261412 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGACTGCCATGGCGTTAGCT 261471

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
Db     261472 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 261531

Qy        118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
Db     261532 TTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 261591

Qy        178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
Db     261592 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 261651

Qy        238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 297
Db     261652 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 261711

Qy        298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 357
Db     261712 CTGCATACCGGAGACTTTCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 261771

Qy        358 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGC 417
Db     261772 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC 261831

Qy        418 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 477
Db     261832 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 261891

Qy        478 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 537
Db     261892 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 261951

Qy        538 AAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 597
Db     261952 AAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 262011

Qy        598 GATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGG 657
Db     262012 GATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGG 262071

Qy        658 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 717
Db     262072 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 262131

Qy        718 AATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTC 777
Db     262132 AATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTC 262191

Qy        778 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCA 837
Db     262192 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCA 262251

Qy        838 GGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 897
Db     262252 GGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 262311

Qy        898 GCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 957
Db     262312 GACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 262371

Qy        958 GGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAAT 1017
Db     262372 GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAAT 262431

Qy       1018 GACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAA 1077
Db     262432 GACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAA 262491

Qy       1078 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA----------- 1126
Db     262492 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATG 262551

Qy       1127 ------------------------------ATGGCTCCAACCATTGCGGAATTACTGCC 1155
Db     262552 TATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCC 262611

Qy       1156 ATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 1215
Db     262612 ATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 262671

Qy       1216 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 1275
Db     262672 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 262731

Qy       1276 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAG 1335
Db     262732 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAG 262791

Qy       1336 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 1395
Db     262792 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 262851

Qy       1396 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 1455
Db     262852 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 262911

Qy       1456 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 1515
Db     262912 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 262971

Qy       1516 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 1575
Db     262972 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 263031

Qy       1576 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 1635
Db     263032 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 263091

Qy       1636 GCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTAT 1695
Db     263092 GCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT 263151

Qy       1696 CTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAG 1755
Db     263152 CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAG 263211

Qy       1756 CTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 1815
Db     263212 CTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 263271

Qy       1816 GCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAAC 1875
Db     263272 GCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAAC 263331

Qy       1876 CGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCGGATTATTCTTTC 1935
```

```
Db 263332 AGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTC 263391

Qy   1936 TTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 1995
          ||||||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db 263392 TTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 263451

Qy   1996 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 2055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263452 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 263511

Qy   2056 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 2115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263512 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 263571

Qy   2116 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 2175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263572 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 263631

Qy   2176 ACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 2235
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 263632 ACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 263691

Qy   2236 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 2295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263692 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 263751

Qy   2296 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 2355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263752 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 263811

Qy   2356 CGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATC 2415
          ||||| ||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db 263812 CGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATC 263871

Qy   2416 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 2475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263872 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 263931

Qy   2476 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 2535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263932 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 263991

Qy   2536 ACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 2595
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db 263992 ACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 264051

Qy   2596 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 2655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264052 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 264111

Qy   2656 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 2715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264112 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 264171

Qy   2716 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 2775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264172 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 264231

Qy   2776 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 2835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264232 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 264291

Qy   2836 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 2895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264292 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 264351

Qy   2896 ACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 2955
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264352 ACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 264411

Qy   2956 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 3015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264412 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 264471

Qy   3016 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGC 3075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db 264472 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGT 264531

Qy   3076 CTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 3135
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db 264532 CTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 264591

Qy   3136 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 3195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264592 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 264651

Qy   3196 CGGTTCTGA 3204
          |||||||||
Db 264652 CGGTTCTGA 264660
```

```
RESULT 4
AAS43905
ID   AAS43905 standard; DNA; 4218 BP.
XX
AC   AAS43905;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein 961cL-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27608.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 23; Page 75-76; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;

  Query Match            93.2%;  Score 2986.2;  DB 22;  Length 4218;
  Best Local Similarity  97.1%;  Pred. No. 0;
  Matches 3067;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;
```

```
Qy      90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
           |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    1056 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1115

Qy     150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1116 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1175

Qy     210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1176 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1235

Qy     270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    1236 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1295

Qy     330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1296 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1355

Qy     390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    1356 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1415

Qy     450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1416 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1475

Qy     510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    1476 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1535

Qy     570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1536 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1595

Qy     630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
           |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1596 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1655

Qy     690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
           |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||| ||
Db    1656 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1715

Qy     750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
           |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1716 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1775

Qy     810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
           ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    1776 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1835

Qy     870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
           ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    1836 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1895

Qy     930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
           |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    1896 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1955

Qy     990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
           |||||||||||| |||||| || |||||||||||||||||||||||||||||    ||||
Db    1956 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 2015

Qy    1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    2016 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2075

Qy    1110 CAGTGGAGAAAAGTTCA------------------------------------------A 1127
           |||||||||||||||||                                          |
Db    2076 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2135

Qy    1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    2136 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2195

Qy    1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2196 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2255

Qy    1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2256 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2315

Qy    1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
           |||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||||
Db    2316 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2375

Qy    1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2376 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2435

Qy    1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2436 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2495

Qy    1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2496 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2555

Qy    1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2556 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2615

Qy    1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2616 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2675

Qy    1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
            ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    2676 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2735

Qy    1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
           |||||| ||||||||||||  ||||||||||| |||| ||||||||||||||||||||||
Db    2736 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2795

Qy    1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
           ||||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db    2796 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2855

Qy    1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
           ||||||||| |||||||||||||||||| ||||||||| |||||||||||||||||||||
Db    2856 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2915

Qy    1908 CGCCAAAATCGGGCGGGATTATTCTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
           |||||||||||||| ||||||||||||||||||||||||||||||||||| || ||||||
Db    2916 CGCCAAAATCGGGCAGGATTATTCTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2975

Qy    1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2976 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 3035

Qy    2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3036 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3095

Qy    2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3096 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3155

Qy    2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    3156 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3215

Qy    2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3216 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3275

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3276 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3335

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   3336 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3395

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3396 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3455

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3456 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3515

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   3516 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3575

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3576 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3635

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3636 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3695

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3696 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3755

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3756 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3815

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3816 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3875

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   3876 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3935

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3936 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3995

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3996 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 4055

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
          ||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
Db   4056 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4115

Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4116 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4175

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          |||||||||||||||||||||||||||||||||||||
Db   4176 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212


RESULT 5
AAD17058
ID   AAD17058 standard; DNA; 4218 BP.
XX
AC   AAD17058;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 961cL-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   961cL-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4212
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 961cL-983
FT                   fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10043.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 23; Page 33-34; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) 961cL-983
CC   fusion protein.
XX
SQ   Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;

  Query Match             93.2%;  Score 2986.2;  DB 22;  Length 4218;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3067;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy     90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
          |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   1056 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1115

Qy    150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1116 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1175

Qy    210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1176 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1235

Qy    270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   1236 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1295

Qy    330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1296 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1355
```

```
Qy    390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
Db   1356 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1415

Qy    450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
Db   1416 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1475

Qy    510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
Db   1476 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1535

Qy    570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
Db   1536 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1595

Qy    630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
Db   1596 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1655

Qy    690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
Db   1656 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1715

Qy    750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
Db   1716 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1775

Qy    810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
Db   1776 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1835

Qy    870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
Db   1836 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1895

Qy    930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
Db   1896 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1955

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
Db   1956 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 2015

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
Db   2016 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2075

Qy   1110 CAGTGGAGAAAAGTTCA-----------------------------------------A 1127
Db   2076 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2135

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
Db   2136 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2195

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
Db   2196 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2255

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
Db   2256 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2315

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
Db   2316 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2375

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
Db   2376 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2435

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
Db   2436 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2495

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
Db   2496 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2555

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
Db   2556 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2615

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
Db   2616 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2675

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
Db   2676 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2735

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
Db   2736 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2795

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
Db   2796 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2855

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
Db   2856 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2915

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
Db   2916 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2975

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
Db   2976 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 3035

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
Db   3036 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3095

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
Db   3096 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3155

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
Db   3156 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3215

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
Db   3216 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3275

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
Db   3276 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3335

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
Db   3336 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3395

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
Db   3396 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3455

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
Db   3456 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3515

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
```

```
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     3516 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3575

Qy     2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3576 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3635

Qy     2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3636 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3695

Qy     2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3696 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3755

Qy     2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3756 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3815

Qy     2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3816 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3875

Qy     2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
            |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     3876 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3935

Qy     2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3936 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3995

Qy     2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3996 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 4055

Qy     3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
            ||||||||||||||||||||||||||| |||||||||||| |||||||||||||||||||
Db     4056 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4115

Qy     3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4116 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4175

Qy     3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
            |||||||||||||||||||||||||||||||||||||
Db     4176 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212
```

```
RESULT 6
AAS43899
ID   AAS43899 standard; DNA; 4335 BP.
XX
AC   AAS43899;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein 961-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27602.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 23; Page 68-69; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;

  Query Match             93.1%;  Score 2983.8;  DB 22;  Length 4335;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches   57;  Indels   42;  Gaps    1;

Qy       75 TTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCAC 134
            ||  |  ||    |  |||||| ||||||||||||||||||||||||||||||||| ||
Db     1140 TTACGAGTGGGGATCCGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTAC 1199

Qy      135 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 194
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1200 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 1259

Qy      195 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 254
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1260 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 1319

Qy      255 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 314
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1320 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 1379

Qy      315 TACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 374
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1380 TCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 1439

Qy      375 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCAT 434
            |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     1440 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCAT 1499

Qy      435 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 494
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1500 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 1559

Qy      495 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTC 554
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     1560 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTC 1619

Qy      555 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 614
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1620 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 1679

Qy      615 AGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 674
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db     1680 AGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 1739

Qy      675 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAAC 734
            |||||||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db     1740 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAAC 1799
```

```
Qy     735 CAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 794
           |||||||||||| |||  ||||||||||||||||||||||||||||||||||||||||||
Db    1800 CAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 1859

Qy     795 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTT 854
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1860 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTT 1919

Qy     855 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGA 914
           |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    1920 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGA 1979

Qy     915 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCA 974
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    1980 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCA 2039

Qy     975 CATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCC 1034
           ||||||||||||||||||||||||||| |||||| || ||||||||||||||||||||||
Db    2040 CATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCC 2099

Qy    1035 CAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGT 1094
           |||||||    ||||||||||||||||||||||||| |||||||||||||||||||||||
Db    2100 CAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGT 2159

Qy    1095 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA---------------------------- 1126
           ||||||||||||||||||||||||||||||||
Db    2160 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTAC 2219

Qy    1127 --------------ATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGC 1172
                         |||||||||||||||||||||||||||||||||||||||| |||||
Db    2220 AGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGC 2279

Qy    1173 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 1232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2280 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 2339

Qy    1233 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 1292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2340 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 2399

Qy    1293 GAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGG 1352
           ||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||
Db    2400 GAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGG 2459

Qy    1353 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 1412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2460 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 2519

Qy    1413 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 1472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2520 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 2579

Qy    1473 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 1532
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2580 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 2639

Qy    1533 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 1592
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2640 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 2699

Qy    1593 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 1652
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2700 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 2759

Qy    1653 GGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCG 1712
           ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    2760 GGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCA 2819

Qy    1713 ATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTAC 1772
           ||||||||||||||||||||| ||||||||||||  ||||||||||| |||| |||||||
Db    2820 ATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTAC 2879

Qy    1773 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAA 1832
           |||||||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db    2880 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAA 2939

Qy    1833 GCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGT 1892
           |||||||||||||||||||||||| |||||||||||||||||| ||||||||| ||||||
Db    2940 GCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGT 2999

Qy    1893 TCCCTTCCTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAAC 1952
           ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    3000 TCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAAC 3059

Qy    1953 CGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 2012
           |||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3060 CGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 3119

Qy    2013 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 2072
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3120 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 3179

Qy    2073 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 2132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3180 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 3239

Qy    2133 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 2192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3240 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 3299

Qy    2193 CGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 2252
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3300 AGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 3359

Qy    2253 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 2312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3360 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 3419

Qy    2313 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGT 2372
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    3420 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGT 3479

Qy    2373 ATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 2432
           |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db    3480 ATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 3539

Qy    2433 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 2492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3540 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 3599

Qy    2493 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACA 2552
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3600 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACG 3659

Qy    2553 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 2612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3660 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 3719

Qy    2613 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 2672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3720 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 3779

Qy    2673 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 2732
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3780 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 3839

Qy    2733 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 2792
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3840 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 3899

Qy    2793 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 2852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3900 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 3959

Qy    2853 AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGC 2912
```

```
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     3960 AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGC 4019

Qy     2913 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 2972
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4020 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 4079

Qy     2973 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 3032
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4080 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 4139

Qy     3033 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGG 3092
            ||||||||||||||||||||||||||||||||||||||||| ||||||||||| |||||
Db     4140 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGG 4199

Qy     3093 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 3152
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4200 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 4259

Qy     3153 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db     4260 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308
```

```
RESULT 7
AAD17052
ID   AAD17052 standard; DNA; 4335 BP.
XX
AC   AAD17052;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 961-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein; 961-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4335
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 961-983
FT                   fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10037.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 23; Page 28-29; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) 961-983 fusion
CC   protein.
XX
SQ   Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;

  Query Match             93.1%;  Score 2983.8;  DB 22;  Length 4335;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches   57;  Indels   42;  Gaps    1;

Qy       75 TTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCAC 134
            ||  |  ||    | |||||| |||||||||||||||||||||||||||||||||| ||
Db     1140 TTACGAGTGGGGATCCGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTAC 1199

Qy      135 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 194
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1200 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 1259

Qy      195 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 254
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1260 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 1319

Qy      255 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 314
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1320 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 1379

Qy      315 TACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 374
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1380 TCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 1439

Qy      375 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCAT 434
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     1440 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCAT 1499

Qy      435 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 494
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1500 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 1559

Qy      495 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTC 554
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     1560 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTC 1619

Qy      555 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 614
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1620 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 1679

Qy      615 AGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 674
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db     1680 AGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 1739

Qy      675 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAAC 734
            ||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db     1740 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAAC 1799

Qy      735 CAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 794
            ||||||||||||  |||  |||||||||||||||||||||||||||||||||||||||||
Db     1800 CAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 1859

Qy      795 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTT 854
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     1860 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTT 1919

Qy      855 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGA 914
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     1920 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGA 1979

Qy      915 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCA 974
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     1980 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCA 2039

Qy      975 CATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCC 1034
            ||||||||||||||||||||||||||| ||||| || |||||||||||||||||||||||
Db     2040 CATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCC 2099
```

```
Qy   1035 CAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGT 1094
          |||||||    ||||||||||||||||||||||||| |||||||||||||||||||||||
Db   2100 CAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGT 2159

Qy   1095 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA---------------------------- 1126
          ||||||||||||||||||||||||||||||||
Db   2160 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTAC 2219

Qy   1127 --------------ATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGC 1172
                        |||||||||||||||||||||||||||||||||||||||| |||||
Db   2220 AGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGC 2279

Qy   1173 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 2339

Qy   1233 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 2399

Qy   1293 GAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGG 1352
          ||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||
Db   2400 GAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGG 2459

Qy   1353 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 2519

Qy   1413 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 2579

Qy   1473 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 2639

Qy   1533 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2640 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 2699

Qy   1593 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 2759

Qy   1653 GGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCG 1712
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   2760 GGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCA 2819

Qy   1713 ATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTAC 1772
          ||||||||||||||||||||| ||||||||||||  ||||||||||| |||| |||||||
Db   2820 ATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTAC 2879

Qy   1773 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAA 1832
          |||||||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db   2880 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAA 2939

Qy   1833 GCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGT 1892
          |||||||||||||||||||||||| |||||||||||||||||| ||||||||| ||||||
Db   2940 GCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGT 2999

Qy   1893 TCCCTTCCTGAGTGCCGCCAAAATCGGGCGGATTATTCTTTCTTCACAAACATCGAAAC 1952
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   3000 TCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAAC 3059

Qy   1953 CGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 2012
          |||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3060 CGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 3119

Qy   2013 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 2072
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3120 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 3179

Qy   2073 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 2132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3180 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 3239

Qy   2133 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 2192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3240 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 3299

Qy   2193 CGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 2252
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3300 AGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 3359

Qy   2253 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 2312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3360 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 3419

Qy   2313 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGT 2372
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   3420 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGT 3479

Qy   2373 ATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 2432
          |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db   3480 ATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 3539

Qy   2433 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 2492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3540 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 3599

Qy   2493 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACA 2552
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3600 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACG 3659

Qy   2553 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 2612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3660 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 3719

Qy   2613 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 2672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3720 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 3779

Qy   2673 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 2732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3780 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 3839

Qy   2733 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 2792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3840 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 3899

Qy   2793 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 2852
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3900 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 3959

Qy   2853 AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGC 2912
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   3960 AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGC 4019

Qy   2913 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 2972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4020 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 4079

Qy   2973 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 3032
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4080 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 4139

Qy   3033 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGG 3092
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||| |||||
Db   4140 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGG 4199

Qy   3093 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 3152
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4200 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 4259

Qy   3153 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
```

```
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db      4260 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308

RESULT 8
AAS43880
ID   AAS43880 standard; DNA; 3939 BP.
XX
AC   AAS43880;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G741-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27581.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 51-52; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 3939 BP; 1046 A; 1103 C; 1106 G; 684 T; 0 other;

  Query Match             93.1%;  Score 2983.2;  DB 22;  Length 3939;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy        90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
             |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       759 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 818

Qy       150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       819 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 878

Qy       210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       879 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 938

Qy       270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
             |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db       939 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 998

Qy       330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       999 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1058

Qy       390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1059 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1118

Qy       450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1119 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1178

Qy       510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      1179 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1238

Qy       570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1239 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1298

Qy       630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1299 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1358

Qy       690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
             |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||| ||
Db      1359 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1418

Qy       750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
             |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1419 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1478

Qy       810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
             ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      1479 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1538

Qy       870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
             ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      1539 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1598

Qy       930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db      1599 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1658

Qy       990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
             |||||||||||| |||||| || |||||||||||||||||||||||||||||    ||||
Db      1659 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1718

Qy      1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1719 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1778

Qy      1110 CAGTGGAGAAAAGTTCA------------------------------------------A 1127
             |||||||||||||||||                                          |
Db      1779 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 1838

Qy      1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      1839 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 1898

Qy      1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1899 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1958

Qy      1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1959 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2018

Qy      1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
             |||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||||
Db      2019 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2078
```

```
Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2079 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2138

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2139 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2198

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2199 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2258

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2259 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2318

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2319 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2378

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   2379 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2438

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
          |||||| ||||||||||||  ||||||||||| |||| ||||||||||||||||||||||
Db   2439 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2498

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
          ||||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db   2499 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2558

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
          ||||||||| |||||||||||||||||| ||||||||| |||||||||||||||||||||
Db   2559 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2618

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
          |||||||||||||| |||||||||||||||||||||||||||||||||||| || |||||
Db   2619 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2678

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2738

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2798

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2858

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   2859 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 2918

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2919 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2978

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3038

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   3039 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3098

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3099 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3158

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3218

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   3219 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3278

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3338

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3339 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3398

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3399 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3458

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3459 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3518

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3519 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3578

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   3579 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3638

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3639 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3698

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3699 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3758

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
          ||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
Db   3759 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 3818

Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3819 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3878

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          ||||||||||||||||||||||||||||||||||
Db   3879 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912
```

```
RESULT 9
AAD17045
ID   AAD17045 standard; DNA; 3939 BP.
XX
AC   AAD17045;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G741-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G741-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3939
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G741-983 fusion protein"
```

```
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10030.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 4; Page 20-21; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G741-983
CC   fusion protein.
XX
SQ   Sequence 3939 BP; 1046 A; 1103 C; 1106 G; 684 T; 0 other;

  Query Match             93.1%;  Score 2983.2;  DB 22;  Length 3939;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy         90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
              |||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        759 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 818

Qy        150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 878

Qy        210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 938

Qy        270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        939 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 998

Qy        330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1058

Qy        390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1059 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1118

Qy        450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1178

Qy        510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       1179 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1238

Qy        570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1239 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1298

Qy        630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1299 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1358

Qy        690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
              |||||||||||||||||||||||||||||||||||| |||||| |||||||||||||| ||
Db       1359 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1418

Qy        750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
              |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1419 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1478

Qy        810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1479 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1538

Qy        870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1539 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1598

Qy        930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1599 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1658

Qy        990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
              ||||||||||| |||||| || ||||||||||||||||||||||||||||     ||||
Db       1659 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1718

Qy       1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1719 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1778

Qy       1110 CAGTGGAGAAAAGTTCA-----------------------------------------A 1127
              |||||||||||||||||                                         |
Db       1779 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 1838

Qy       1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       1839 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 1898

Qy       1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1899 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1958

Qy       1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1959 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2018

Qy       1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
              |||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||||
Db       2019 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2078

Qy       1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2079 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2138

Qy       1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2139 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2198

Qy       1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2199 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2258

Qy       1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2259 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2318

Qy       1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2319 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2378

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   2379 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2438

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
          |||||| ||||||||||||  ||||||||||| |||| ||||||||||||||||||||||
Db   2439 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2498

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
          ||||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db   2499 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2558

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
          ||||||||| |||||||||||||||||| ||||||||| |||||||||||||||||||||
Db   2559 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2618

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
          |||||||||||||| |||||||||||||||||||||||||||||||||||| || |||||
Db   2619 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2678

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2738

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2798

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2858

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   2859 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 2918

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2919 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2978

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3038

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   3039 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3098

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3099 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3158

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3218

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   3219 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3278

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3338

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3339 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3398

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3399 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3458

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3459 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3518

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3519 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3578

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   3579 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3638

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGCGTGGAACGCGACCTGAA 2987
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3639 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGCGTGGAACGCGACCTGAA 3698

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3699 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3758

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
          ||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
Db   3759 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 3818

Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3819 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3878

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          ||||||||||||||||||||||||||||||||||
Db   3879 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912


RESULT 10
AAS43902
ID   AAS43902 standard; DNA; 4170 BP.
XX
AC   AAS43902;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein 961c-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27605.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 23; Page 71-73; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
```

```
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4170 BP; 1185 A; 1138 C; 1119 G; 728 T; 0 other;

  Query Match             93.1%;  Score 2983.2;  DB 22;  Length 4170;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy     90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
          |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    990 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1049

Qy    150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1109

Qy    210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1169

Qy    270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   1170 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1229

Qy    330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1289

Qy    390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1290 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1349

Qy    450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1409

Qy    510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1410 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1469

Qy    570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1470 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1529

Qy    630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1589

Qy    690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
          |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||| ||
Db   1590 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1649

Qy    750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
          |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1709

Qy    810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1710 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1769

Qy    870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1770 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1829

Qy    930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1830 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1889

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
          |||||||||||| |||||| || |||||||||||||||||||||||||||||    ||||
Db   1890 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1949

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1950 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2009

Qy   1110 CAGTGGAGAAAAGTTCA------------------------------------------A 1127
          |||||||||||||||||                                          |
Db   2010 CAGTGGAGAAAAGTTCAAACGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2069

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   2070 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2129

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2130 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2189

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2190 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2249

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
          |||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||||
Db   2250 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2309

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2310 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2369

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2370 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2429

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2430 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2489

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2490 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2549

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2550 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2609

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   2610 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2669

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
          |||||| ||||||||||||  ||||||||||| |||| ||||||||||||||||||||||
Db   2670 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2729

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
          ||||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db   2730 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2789

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
          ||||||||| |||||||||||||||||| ||||||||| |||||||||||||||||||||
Db   2790 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2849

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
          |||||||||||||| |||||||||||||||||||||||||||||||||||| || |||||
Db   2850 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2909

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2910 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2969

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2970 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3029

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3030 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3089

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   3090 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3149

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3150 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3209

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3210 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3269

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   3270 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3329

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3330 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3389

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3390 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3449

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   3450 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3509

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3510 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3569

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3570 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3629

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3630 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3689

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3690 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3749

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3750 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3809

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   3810 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3869

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3870 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3929

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3930 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3989

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
          ||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
Db   3990 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4049

Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4050 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4109

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          ||||||||||||||||||||||||||||||||||
Db   4110 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143


RESULT 11
AAD17055
ID   AAD17055 standard; DNA; 4170 BP.
XX
AC   AAD17055;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 961c-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta 961c-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4170
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 961c-983
FT                   fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10040.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 23; Page 31; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) 961c-983 fusion
CC   protein.
XX
SQ   Sequence 4170 BP; 1185 A; 1138 C; 1119 G; 728 T; 0 other;

  Query Match             93.1%;  Score 2983.2;  DB 22;  Length 4170;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy     90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
          |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    990 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1049
```

```
Qy    150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1109

Qy    210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1169

Qy    270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   1170 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1229

Qy    330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1289

Qy    390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1290 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1349

Qy    450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1409

Qy    510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1410 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1469

Qy    570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1470 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1529

Qy    630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1589

Qy    690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
          |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||| ||
Db   1590 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1649

Qy    750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
          |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1709

Qy    810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1710 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1769

Qy    870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1770 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1829

Qy    930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1830 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1889

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
          |||||||||||| |||||| || |||||||||||||||||||||||||||||    ||||
Db   1890 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1949

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1950 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2009

Qy   1110 CAGTGGAGAAAAGTTCA------------------------------------------A 1127
          |||||||||||||||||                                          |
Db   2010 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2069

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   2070 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2129

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2130 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2189

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2190 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2249

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
          |||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||||
Db   2250 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2309

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2310 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2369

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2370 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2429

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2430 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2489

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2490 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2549

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2550 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2609

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   2610 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2669

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
          |||||| ||||||||||||  ||||||||||| |||| ||||||||||||||||||||||
Db   2670 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2729

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
          ||||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db   2730 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2789

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
          ||||||||| |||||||||||||||||| ||||||||| |||||||||||||||||||||
Db   2790 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2849

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
          |||||||||||||| |||||||||||||||||||||||||||||||||||| || |||||
Db   2850 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2909

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2910 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2969

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2970 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3029

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3030 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3089

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   3090 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3149

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3150 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3209

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3210 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3269

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   3270 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3329

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3330 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3389

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3390 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3449

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   3450 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3509

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3510 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3569

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3570 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3629

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3630 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3689

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3690 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3749

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3750 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3809

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   3810 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3869

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3870 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3929

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3930 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3989

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
          ||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
Db   3990 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4049

Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4050 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4109

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          ||||||||||||||||||||||||||||||||||
Db   4110 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143


RESULT 12
AAS43875
ID   AAS43875 standard; DNA; 3939 BP.
XX
AC   AAS43875;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-741 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27576.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 44-45; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 3939 BP; 1046 A; 1100 C; 1107 G; 686 T; 0 other;

  Query Match             92.8%;  Score 2971.8;  DB 22;  Length 3939;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy    103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy    163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy    223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy    283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy    343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy    403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy    463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423
```

```
Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
Db    424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
Db    484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy    643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
Db    544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy    703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
Db    604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy    763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
Db    664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy    823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
Db    724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy    883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
Db    784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy    943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
Db    844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy   1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
Db    904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy   1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
Db    964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy   1123 TTCA----------------------------------------ATGGCTCCAACCAT 1140
Db   1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy      2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy      2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy      2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy      2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
             ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy      2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy      3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy      3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
             |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db      3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy      3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy      3181 GTCGGCGTAGGCTACCGGTTC 3201
             |||||||||||||||||||||
Db      3124 GTCGGCGTAGGCTACCGGTTC 3144
```

RESULT 13
AAD17040
ID   AAD17040 standard; DNA; 3939 BP.
XX
AC   AAD17040;
XX
DT   29-NOV-2001 (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-741 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G983-741 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..3939
FT   /*tag= a
FT   /product= "N. meningitidis strain 2996 delta
FT   G983-741 fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10023.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state -
XX
PS   Example 3; Page 15-16; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-741
CC   fusion protein.
XX
SQ   Sequence 3939 BP; 1046 A; 1100 C; 1107 G; 686 T; 0 other;

```
  Query Match             92.8%;  Score 2971.8;  DB 22;  Length 3939;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy       103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
             ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy       163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy       223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy       283 GCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       184 GCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy       343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy       403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy       463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy       523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy       583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy       643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy       703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
             ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db       604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy       763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy    823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
Db    724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy    883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
Db    784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy    943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
Db    844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy   1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
Db    904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy   1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
Db    964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy   1123 TTCA----------------------------------------ATGGCTCCAACCAT 1140
Db   1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883
```

```
Qy     2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy     3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGCACGCAAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGCACGCAAT 3003

Qy     3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
            |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db     3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy     3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy     3181 GTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||
Db     3124 GTCGGCGTAGGCTACCGGTTC 3144
```

RESULT 14
AAS43877

```
ID   AAS43877 standard; DNA; 4179 BP.
XX
AC   AAS43877;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-961c DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27578.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 47-49; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;

  Query Match             92.8%;  Score 2971.8;  DB 22;  Length 4179;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy      103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy      163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy      223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy      283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
            ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy      343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy      403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy      463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy      523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy      583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy      643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy      703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
            ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db      604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy      763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy      823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy      883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy      943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
            ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db      844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy     1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
            |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db      904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy     1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy     1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
            ||||                                          ||||||||||||||
```

```
Db   1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
          ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
          |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
          ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db   3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy   3181 GTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||
Db   3124 GTCGGCGTAGGCTACCGGTTC 3144
```

RESULT 15
AAD17042
ID   AAD17042 standard; DNA; 4179 BP.

```
XX
AC   AAD17042;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-961c fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G983-961c fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4179
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G983-961c fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10025.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 3; Page 17-18; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-961c
CC   fusion protein.
XX
SQ   Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;

  Query Match            92.8%;  Score 2971.8;  DB 22;  Length 4179;
  Best Local Similarity  97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy        103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy        163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy        223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy        283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy        343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy        403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy        463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy        523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy        583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy        643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy        703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
              ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db        604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy        763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy        823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy        883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy        943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db        844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy       1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
              |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db        904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy       1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy       1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
              ||||                                          ||||||||||||||
Db       1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy    1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy    1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy    1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy    1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
                ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy    1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
                ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db    1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy    1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
                ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db    1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy    1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
                |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db    1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy    1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
                ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db    1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy    1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
                | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db    1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy    1981 AGCGTCGAAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1924 AGCGTCGAAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy    2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy    2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy    2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
                |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy    2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy    2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy    2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db    2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy    2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy    2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy    2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
                ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy    2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy    2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy    2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy    2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy    2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy    2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
                ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy    2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy    3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy    3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
                |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db    3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy    3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy    3181 GTCGGCGTAGGCTACCGGTTC 3201
                |||||||||||||||||||||
Db    3124 GTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 16
AAS43876
ID   AAS43876 standard; DNA; 4344 BP.
XX
AC   AAS43876;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-961 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
```

```
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27577.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 46-47; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4344 BP; 1219 A; 1191 C; 1167 G; 767 T; 0 other;

  Query Match             92.8%;  Score 2971.8;  DB 22;  Length 4344;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy        103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy        163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy        223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy        283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy        343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy        403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy        463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy        523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy        583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy        643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy        703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
              ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db        604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy        763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy        823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy        883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy        943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db        844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy       1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
              |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db        904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy       1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy       1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
              ||||                                          ||||||||||||||
Db       1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy       1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy       1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db       1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy       1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
```

```
          ||||||  |||||||||| |||| ||||||||||||||||||||||||||||||||||||
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
          |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
          ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db   3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy   3181 GTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||
Db   3124 GTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 17
AAS43874
ID   AAS43874 standard; DNA; 4425 BP.
XX
AC   AAS43874;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-ORF46.1 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27575.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 42-43; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
```

```
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;

  Query Match             92.8%;  Score 2971.8;  DB 22;  Length 4425;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy        103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy        163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy        223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy        283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy        343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy        403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy        463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy        523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy        583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy        643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy        703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
              ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db        604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy        763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy        823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy        883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy        943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db        844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy       1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
              |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db        904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy       1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy       1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
              ||||                                          ||||||||||||||
Db       1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy       1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy       1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db       1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy       1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
              ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db       1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy       1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
              |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db       1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy       1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
              ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db       1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy       1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
              | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db       1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy       1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy       2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy       2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db   3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy   3181 GTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||
Db   3124 GTCGGCGTAGGCTACCGGTTC 3144

RESULT 18
AAD17039
ID   AAD17039 standard; DNA; 4425 BP.
XX
AC   AAD17039;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-ORF46.1 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein; open reading frame; ORF;
KW   delta G983-ORF46.1 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4425
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G983-ORF46.1 fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10022.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 3; Page 13-14; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-ORF46.1
CC   (open reading frame) fusion protein.
XX
SQ   Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;

  Query Match             92.8%;  Score 2971.8;  DB 22;  Length 4425;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy    103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy    163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy    223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy    283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
Db    184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy    343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
Db    244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy    403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
Db    304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy    463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
Db    364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
Db    424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
Db    484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy    643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
Db    544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy    703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
Db    604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy    763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
Db    664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy    823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
Db    724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy    883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
Db    784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy    943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
Db    844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy   1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
Db    904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy   1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
Db    964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy   1123 TTCA-----------------------------------------ATGGCTCCAACCAT 1140
Db   1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343
```

```
Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db   3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy   3181 GTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||
Db   3124 GTCGGCGTAGGCTACCGGTTC 3144


RESULT 19
AAD17041
ID   AAD17041 standard; DNA; 3798 BP.
XX
AC   AAD17041;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-961 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G983-961 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3798
FT                   /*tag= a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G983-961 fusion protein"
FT                   /transl_except= (pos:2104..2109, aa:Xaa)
FT                   /note= "Xaa corresponds to residues 702-885 of delta
FT                   G983-961 fusion protein; There is apparent deletion
FT                   of codons corresponding to residues 702-885 of this
FT                   protein, which alters the reading frame"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10024.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 3; Page 16-17; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-961
CC   fusion protein.
XX
SQ   Sequence 3798 BP; 1079 A; 1020 C; 1007 G; 692 T; 0 other;

  Query Match             60.8%;  Score 1946.6;  DB 22;  Length 3798;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches 2022;  Conservative    0;  Mismatches   39;  Indels   42;  Gaps    1;

Qy    103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy    163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy    223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy    283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy    343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy    403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363
```

```
Qy      463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy      523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy      583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy      643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy      703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
            ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db      604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy      763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy      823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy      883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy      943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
            ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db      844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy     1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
            |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db      904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy     1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy     1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
            ||||                                          ||||||||||||||
Db     1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy     1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
            |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy     1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy     1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy     1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
            ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy     1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy     1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy     1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy     1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy     1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
            ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy     1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
            ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db     1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy     1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
            ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db     1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy     1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
            |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db     1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy     1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
            ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db     1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy     1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
            | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db     1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy     1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy     2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy     2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy     2161 TCC 2163
            |||
Db     2104 TCC 2106
```

```
RESULT 20
AAZ53301
ID   AAZ53301 standard; DNA; 1365 BP.
XX
AC   AAZ53301;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 140 partial DNA sequence SEQ ID NO:551.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
```

```
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74539.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 396-397; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1365 BP; 319 A; 410 C; 418 G; 218 T; 0 other;

  Query Match             42.1%;  Score 1349;  DB 21;  Length 1365;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1355;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1840 ATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTC 1899
          ||||||||| |||| || ||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTCGGCAGGCGGTAAGGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTC 60

Qy   1900 CTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGT 1959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db     61 CTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGC 120

Qy   1960 GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTG 2019
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    121 GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGTAGTGAAGGCGACACGCTG 180

Qy   2020 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 2079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 240

Qy   2080 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 2139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 300

Qy   2140 GAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGAC 2199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGAC 360

Qy   2200 CGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTA 2259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTA 420

Qy   2260 CAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT 2319
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    421 CAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAATCTCGCCGCTACCGTCTAT 480

Qy   2320 GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGAC 2379
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    481 GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGAC 540

Qy   2380 GGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGA 2439
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGA 600

Qy   2440 ACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATT 2499
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATT 660

Qy   2500 GCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACA 2559
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACA 720

Qy   2560 TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGG 2619
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGG 780

Qy   2620 CACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAAC 2679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAAC 840

Qy   2680 AGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTG 2739
          ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db    841 AGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGCGAGCGTCAACGGCACGCTG 900

Qy   2740 ATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACG 2799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACG 960

Qy   2800 GTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGT 2859
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGT 1020

Qy   2860 GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTG 2919
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   1021 GCTTTGGGCTGGAGCGGCAACAGCATCACTGAAGGCACACTGGTCGGACTCGCGGGTCTG 1080

Qy   2920 AAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGC 2979
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGC 1140

Qy   2980 GACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACC 3039
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACC 1200

Qy   3040 GGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGAT 3099
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGAT 1260

Qy   3100 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAG 3159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAG 1320

Qy   3160 TACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 1365


RESULT 21
AAZ53300
ID   AAZ53300 standard; DNA; 1365 BP.
XX
AC   AAZ53300;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 140 partial DNA sequence SEQ ID NO:549.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
```

```
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74538.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 395; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1365 BP; 317 A; 413 C; 421 G; 214 T; 0 other;

  Query Match             41.9%;  Score 1342.6;  DB 21;  Length 1365;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches 1351;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1840 ATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTC 1899
          ||||||||||||||||| |||||||||||||||||| ||||||||| |||||||||||||
Db      1 ATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTC 60

Qy   1900 CTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGT 1959
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     61 CTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGC 120

Qy   1960 GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTG 2019
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTG 180

Qy   2020 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 2079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 240

Qy   2080 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 2139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 300

Qy   2140 GAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGAC 2199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    301 GAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGAC 360

Qy   2200 CGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTA 2259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTA 420

Qy   2260 CAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT 2319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT 480

Qy   2320 GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGAC 2379
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    481 GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGAC 540

Qy   2380 GGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGA 2439
          ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    541 GGGTTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGA 600

Qy   2440 ACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATT 2499
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATT 660

Qy   2500 GCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACA 2559
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    661 GCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACA 720

Qy   2560 TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGG 2619
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGG 780

Qy   2620 CACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAAC 2679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAAC 840

Qy   2680 AGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTG 2739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTG 900

Qy   2740 ATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACG 2799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACG 960

Qy   2800 GTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGT 2859
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGT 1020

Qy   2860 GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTG 2919
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   1021 GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTG 1080

Qy   2920 AAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGC 2979
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGC 1140

Qy   2980 GACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACC 3039
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACC 1200

Qy   3040 GGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGAT 3099
          ||||||||||||||||||||||||||||||||||| ||||||||||| ||||||||||||
Db   1201 GGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGAT 1260

Qy   3100 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAG 3159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAG 1320

Qy   3160 TACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 1365
```

```
RESULT 22
AAZ53299
ID   AAZ53299 standard; DNA; 1365 BP.
XX
AC   AAZ53299;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria gonorrheae ORF 140 partial DNA sequence SEQ ID NO:547.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria gonorrheae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74537.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 394; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1365 BP; 324 A; 405 C; 416 G; 220 T; 0 other;

  Query Match             40.2%;  Score 1288.2;  DB 21;  Length 1365;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches 1317;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy   1840 ATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTC 1899
Db      1 ATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACATGTTCCCTTC 60

Qy   1900 CTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGT 1959
Db     61 CTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCAAAAATATCAAAACCGACGGC 120

Qy   1960 GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTG 2019
Db    121 GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCCG 180

Qy   2020 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 2079
Db    181 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 240

Qy   2080 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 2139
Db    241 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 300

Qy   2140 GAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGAC 2199
Db    301 GAGCTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGTCGCCGAC 360

Qy   2200 CGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTA 2259
Db    361 CGCACAGATATGCCGGGCATCCGCCTACGGCGCACAACTTTCCGCACAGCGGCAGCCGTA 420

Qy   2260 CAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT 2319
Db    421 CAGCATGCGAATACCGCCGACGGCGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT 480

Qy   2320 GCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGAC 2379
Db    481 GCCGACAGTGCCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGAC 540

Qy   2380 GGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGA 2439
Db    541 GGGTTGGACCACAACGGTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGA 600

Qy   2440 ACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATT 2499
Db    601 ACGTGGGAACAGGGCGGTGTCGAAGGCAAAATGCGCGGCAGTACCCAAACTATCGGCATT 660

Qy   2500 GCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACA 2559
Db    661 GCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATAGGACGCAGCACA 720

Qy   2560 TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGG 2619
Db    721 TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGG 780

Qy   2620 CACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAAC 2679
Db    781 CACGATGTGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAAC 840

Qy   2680 AGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTG 2739
Db    841 AGCATCAGCCGCAGCACCGGTGCGGATGAATATGCGGAAGGCAGCGTCAACGGCACGCTG 900

Qy   2740 ATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACG 2799
Db    901 ATGCAGCTGGGCGCACTGGGTGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACG 960

Qy   2800 GTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGT 2859
Db    961 GTTGAAGGCGGTCTGCGCCACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGT 1020

Qy   2860 GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTG 2919
Db   1021 GCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTG 1080

Qy   2920 AAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGC 2979
Db   1081 AAACTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTCTGCGACGGCGGGCGTGGAACGC 1140

Qy   2980 GACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACC 3039
Db   1141 GACCTGAACGGACGCGACTACGCGGTAACGGGCGGCTTTACCGGCGCGGCTGCAGCAACC 1200
```

```
Qy   3040 GGCAAGACGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGAT 3099
          ||||||||||| |||||||||||||||||||||||| |||||||||||||| | ||||
Db   1201 GGCAAGACGGGTGCACGCAATATGCCGCACACCCGCCGGGTTGCCGGTCTGGGGTGGAT 1260

Qy   3100 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAG 3159
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1261 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACACCGGTTCCAAACAG 1320

Qy   3160 TACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||| | |||||||||||||||||||||||
Db   1321 TACGGCAACCACAGCGGACAAATCGGCGTAGGCTACCGGTTCTGA 1365
```

RESULT 23
AAZ53298
ID AAZ53298 standard; DNA; 526 BP.
XX
AC AAZ53298;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 139 partial DNA sequence SEQ ID NO:545.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR P-PSDB; AAY74536.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 7; Page 393; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 526 BP; 159 A; 142 C; 131 G; 94 T; 0 other;

```
  Query Match             15.4%;  Score 494.2;  DB 21;  Length 526;
  Best Local Similarity   98.7%;  Pred. No. 3.1e-127;
  Matches  520;  Conservative    0;  Mismatches    3;  Indels    4;  Gaps    2;

Qy      1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy     61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
          ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db     61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 120

Qy    118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    121 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGGGCAACAACAGCGAAATCAGCA 180

Qy    178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCAATATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 240

Qy    238 GGTCGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 297
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    241 GGTCGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATG-CCCCCCCCCGAAT 299

Qy    298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 359

Qy    358 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGC 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    360 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC 419

Qy    418 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 479

Qy    478 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGA 524
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    480 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGA 526
```

RESULT 24
AAZ53297
ID AAZ53297 standard; DNA; 528 BP.
XX
AC AAZ53297;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 138 partial DNA sequence SEQ ID NO:543.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.

3

```
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74535.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 393; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 528 BP; 159 A; 142 C; 131 G; 96 T; 0 other;

  Query Match             14.7%;  Score 470.6;  DB 21;  Length 528;
  Best Local Similarity   98.1%;  Pred. No. 1.2e-120;
  Matches  519;  Conservative    0;  Mismatches    4;  Indels    6;  Gaps    4;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGACTGCCATGGCGTTAGCT 60

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
              |||||||||||||||||||||||||||||||||||    |||||||||||||||||||||
Db         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 120

Qy        118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 180

Qy        178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 240

Qy        238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCCGAAT 297
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        241 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATG-CCCCCCCCCGAAT 299

Qy        298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCA-TACAAGAATTTGATCAACCTCAA 356
              ||||||||||||||||| |||||||||||||||||| |||||||||||||||||||||||
Db        300 CTGCATACCGGAGACTTTCCAAACCCAAATGACGCATTACAAGAATTTGATCAACCTCAA 359

Qy        357 ACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGG 416
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        360 ACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGG 419

Qy        417 CGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAA 479

Qy        477 CGAAAATTAC-AAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGA 524
              |||||||||| ||||||||||||||||||||||||||||||||||||||
Db        480 CGAAAATTACGAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGA 528
```

```
RESULT 25
AAZ53296
ID   AAZ53296 standard; DNA; 537 BP.
XX
AC   AAZ53296;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria gonorrheae ORF 138 partial DNA sequence SEQ ID NO:541.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria gonorrheae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74534.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 392; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 537 BP; 169 A; 136 C; 130 G; 102 T; 0 other;

  Query Match             13.5%;  Score 433.4;  DB 21;  Length 537;
  Best Local Similarity   93.2%;  Pred. No. 2.8e-110;
  Matches  504;  Conservative    0;  Mismatches   16;  Indels   21;  Gaps    4;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              |||||||| ||| |||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAACCACCTCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
              |||||||||||||||||||||||||||||||||||    |||||||||||||| ||||||
```

```
Db      61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCTCCCGAC 120

Qy     118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
           || ||||||||||||||||||||||||||||||||||| ||||| | |||| ||||||||
Db     121 TTTAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGGGCAACGATAGCGGAATCAGCA 180

Qy     178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
           |||||||||||||||||||| || ||||||||||||||||||||||||||||||||||||
Db     181 GCAGTATCTTACGCCGGTATAAAAAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 240

Qy     238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 297
           |||||||||||||||||||||||||||||||||||||||||||| |    ||||||||||
Db     241 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAAG----CCCCCCGAAT 296

Qy     298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATA-CAAGAATTTGATCAACCTCAA 356
           |||||||||||||||||| ||||||||||||||  |||  |||||| |||||||||||||
Db     297 CTGCATACCGGAGACTTTTCAAACCCAAATGACCAATATTAAGAATATGATCAACCTCAA 356

Qy     357 ACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGG 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     357 ACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGG 416

Qy     417 CGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAA 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     417 CGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAA 476

Qy     477 CGA------------AAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTG 523
           |||            |||||||||||||||||||||||||||||||||||||||||||||
Db     477 CGAAAATTACAAAAACAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTG 536

Qy     524 A 524
           |
Db     537 A 537
```

RESULT 26
AAV03603
ID   AAV03603 standard; DNA; 286 BP.
XX
AC   AAV03603;
XX
DT   22-OCT-1998 (first entry)
XX
DE   Neisseria meningitidis DNA sequence E85.
XX
KW   N. gonorrheae; N. lactamica; chromosome Z2491; region 1; region 2;
KW   region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
KW   meningitis; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9802547-A2.
XX
PD   22-JAN-1998.
XX
PF   11-JUL-1997;   97WO-FR01295.
XX
PR   12-JUL-1996;   96FR-0008768.
XX
PA   (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SMIK) SMITHKLINE BEECHAM.
XX
PI   Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
PI   Vinals C;
XX
DR   WPI; 1998-110594/10.
XX
PT   Genes present in Neisseria meningitidis but not other Neisseria
PT   species - and related host cells, RNA, anti-sense sequences,
PT   polypeptide(s) and antibodies, useful for diagnosing Neisseria
PT   meningitidis infection and in protective vaccines
XX
PS   Example 4; Page 128; 150pp; French.
XX
CC   AAV03575-606 represent sequences that are present in Neisseria
CC   meningitidis and N. gonorrheae but not in N. lactamica, except for the
CC   genes involved in biosynthesis of the capsule polysaccharide, frpA or C,
CC   opc, porA, rotamase, sequence IC1106, IgA protease, pillin, pilC,
CC   proteins which bind transferrin and opacity proteins. The DNA sequences
CC   are responsible for the differences in pathogenicity between N.
CC   meningitidis and N. gonorrheae, specifically they include the genes that
CC   allow N. meningitidis to cross the blood-brain barrier. DNA sequences
CC   common to N. meningitidis and N. gonorrheae, but absent from N.
CC   lactamica, are responsible for colonisation and penetration of the
CC   mucosa. The DNA sequences can be used to produce probes and primers, and
CC   antibodies produced against the encoded proteins are used in standard
CC   hybridisation/immunoassay processes for diagnosis of N. meningitidis
CC   infection, particularly meningitis.
XX
SQ   Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

```
  Query Match             8.5%;  Score 270.8;  DB 19;  Length 286;
  Best Local Similarity   99.0%;  Pred. No. 4e-65;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy     865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy     925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy     985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACAC-ACT 1043
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACAACT 180

Qy    1044 GACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGT 1103
           ||||||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||
Db     181 GACCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTTGCAGGCGT 240

Qy    1104 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 1149
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 286
```

RESULT 27
AAA15396
ID   AAA15396 standard; DNA; 286 BP.
XX
AC   AAA15396;
XX
DT   04-SEP-2000 (first entry)
XX
DE   Genomic fragment of Neisseria meningitidis Z2491.
XX
KW   Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO200026375-A2.
XX
PD   11-MAY-2000.
XX
PF   28-OCT-1999;   99WO-FR02643.
XX
PR   30-OCT-1998;   98FR-0013693.
XX
PA   (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA   (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
PI   Tinsley C, Perrin A;
XX
DR   WPI; 2000-365622/31.

```
XX
PT   New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT   or preventative vaccines and for diagnosis  -
XX
PS   Claim 1; Page 44; 187pp; French.
XX
CC   The present sequence represents a genomic fragment of Neisseria
CC   meningitidis. The specificatino describes proteins that are specific
CC   for pathogenic strains of Neisseria. The polynucleotides, polypeptides,
CC   or their antigenic fragments, are used in vaccines to treat or protect
CC   against Neisseria infections, particularly by N. meningitidis. The
CC   polynucleotide sequences are also used for recombinant production of
CC   the polypeptide and to produce attenuated Neisseria strains that
CC   overexpress it, or express it in a non-toxic mutant form.
CC   note: the present sequence is included in Claim 1, but it is those
CC   sequences that do not include the present sequence that are actually
CC   claimed.
XX
SQ   Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

  Query Match              8.5%;  Score 270.8;  DB 21;  Length 286;
  Best Local Similarity   99.0%;  Pred. No. 4e-65;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy      925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy      985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACAC-ACT 1043
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACAACT 180

Qy     1044 GACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGT 1103
            |||||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||
Db      181 GACCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTTGCAGGCGT 240

Qy     1104 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 1149
            ||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 286


RESULT 28
AAA82003
ID   AAA82003 standard; DNA; 577 BP.
XX
AC   AAA82003;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_550 SEQ ID NO:550.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1652-1653; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 577 BP; 189 A; 134 C; 99 G; 155 T; 0 other;

  Query Match              2.3%;  Score 74.4;  DB 21;  Length 577;
  Best Local Similarity   92.9%;  Pred. No. 3e-10;
  Matches   78;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
            |||  |||||| ||||||||||||||| | |||||||||||| |||||||||||||||||
Db      494 ATGTCAACGACTCCAACCTTCCCTACACAGACTTTCAAACCGACTGCCATGGCGTTAGCT 553

Qy       61 GTTGCAACAACACTTTCTGCCTGC 84
            ||||||||||||||||||||||||
Db      554 GTTGCAACAACACTTTCTGCCTGC 577


RESULT 29
AAA81823/c
ID   AAA81823 standard; DNA; 476 BP.
XX
AC   AAA81823;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_370 SEQ ID NO:370.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
```

```
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1608; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 476 BP; 84 A; 154 C; 125 G; 111 T; 2 other;

  Query Match             2.0%;  Score 65.4;  DB 21;  Length 476;
  Best Local Similarity   70.2%;  Pred. No. 8.8e-08;
  Matches   87;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy     1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
            ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      476 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 417

Qy     1123 TTCAATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAA 1182
            ||||||  |               | ||     ||| |  |  ||    |   |||||   ||
Db      416 TTCAATATCCGTTTGAGCGGTGGCGAGGCTGACGCGCTGTCGGCACTCGCACCNGTTGTC 357

Qy     1183 GCAA 1186
            | ||
Db      356 GGAA 353


RESULT 30
AAA10594
ID   AAA10594 standard; DNA; 10732 BP.
XX
AC   AAA10594;
XX
DT   29-JUN-2000  (first entry)
XX
DE   Gene encoding a subunit of cellulose synthase.
XX
KW   Cellulose synthase; cellulose production; increase yield; ds.
XX
OS   Vigna angularis.
XX
PN   JP2000060568-A.
XX
PD   29-FEB-2000.
XX
PF   26-AUG-1998;   98JP-0239998.
XX
PR   26-AUG-1998;   98JP-0239998.
XX
PA   (MIZU/) MIZUNO K.
PA   (OJIP ) OJI PAPER CO.
XX
DR   WPI; 2000-342371/30.
DR   P-PSDB; AAY85179.
XX
PT   A gene encoding a cellulose synthetic equipment - for the improvement
PT   in the amount of cellulose synthesised in a plant body
XX
PS   Claim 2; Page 14-21; 32pp; Japanese.
XX
CC   This sequence represents a gene encoding a subunit of the cellulose
CC   synthase complex of Vigna angularis. The invention relates to subunits of
CC   cellulose synthetic equipment, that can be used to increase the amount of
CC   cellulose synthesised by a plant. The proteins and genes encoding them
CC   can also be used to improve the properties of the cellulose being
CC   produced by a plant.
XX
SQ   Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;

  Query Match             1.9%;  Score 62;  DB 21;  Length 10732;
  Best Local Similarity   13.9%;  Pred. No. 3.8e-06;
  Matches   98;  Conservative  300;  Mismatches  304;  Indels    3;  Gaps    2;

Qy     2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
              :: :||: : | ||: :||:: : : |:::: :: :  :| :|:|  : | :  : || :
Db     9208 HSGYAAMTSRARGGSRMTGYSASASNTHRSRTHRTYRGVAARGVAGYSGGARGASNGGAR 9267

Qy     2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
              :::  |  ||::|||:  :: ::    |:: ::::: ||::| | :     |  |: |:
Db     9268 GGHSRARGARGTHRAAGVASRHSYSGNGNGSRTHRYSYSAAYSGGCYSGNARGARGVAAR 9327

Qy     2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
              : :| | :|| |    :: :  : |:  :  : :|    : ||    :|:::  : :  :
Db     9328 GYSARGRGYRAAAAAAYSMTYSASNGVAGMTGYARGASSRHGARGARGSRYSSRTHRASN 9387

Qy     2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
             ::  : : :||  :: ::  ::: :::        : |      | :  :    | : ::: :
Db     9388 SRVAGSRSRVAASTHRSRRGTHRRTHRYSASNTHRTHRGGNCYSAAVAGGGASNYSVAYS 9447

Qy     2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
            |:: :   : : : :     :::|:|: :| :       ::  :::::  ::    |::::  :::
Db     9448 GSRASNARGYSASNGGNHSRARGVAMTTHRARGTHRAASRYSHSGSRGNGASHSYSGYHS 9507

Qy     2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
             : :   :::::    :  |:::  : ::   :::     |:::|: ::        : :: ||:
Db     9508 SRVAGHSHSSRASNVAAASRHSGHSRHAASRSRASAAGYSRASASYSTHRSRCYSAAGSR 9567

Qy     2748 GGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
              :  : |  :|   :::: : :: :::::  ::  :  ::::||     :|: : : |  :
Db     9568 TRAASRAASRGA--SHSHARGSRGYSYSYSGRSRCYSYSSRVAGYAASRASASNMTASAS 9625

Qy     2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
             ::  :||::  : :|:      ::|| ::|     :::   :::: |:| ::|       ::  | :|
Db     9626 HVAGMTGYSAAVAVASRVAGYSAATHRGNSRRSRAASRYSAAVATHRGASNSRSRGTHRG 9685

Qy     2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
             :| : : :: ::  | :::      :|: :      | :|::| :       :|: :    :   |:
Db     9686 YTHRYSGSRTHRRGGGYYSGRVARASHSSRGNTHRTHRGTHRTHRRGVAVAASMTGRMTS 9745

Qy     2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
```

```
         :  :::    :: :::: |    :: :: :   : :    :: :|: |:| |:: :|::| :
Db    9746 RASHSSRASSRYSSRYSASNTHRCYSSRASHYSGYTHRVARGY-TRGNASVAVAYSTHRG 9804

Qy    2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
                   ::| ::: ::|::|: : : ::    |   :|::: |:   :: ::|:: ::: :|
Db    9805 GNASNHSVATHRHSYSGYRASASHSASYSVAAAARGHSVAASNASHASHASHASSRSRAR 9864

Qy    3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGG 3092
            || :::|::   ::    :   :  |:: :|:::   :::  : :::
Db    9865 GGYSRVAHSASTHRGNASNRRGNCYSHSCYSSRCYSSRASSRSRV 9909


RESULT 31
AAT91476
ID   AAT91476 standard; DNA; 913 BP.
XX
AC   AAT91476;
XX
DT   07-JAN-1998  (first entry)
XX
DE   Mycobacterium tuberculosis antigen TbRa32 encoding DNA.
XX
KW   Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW   skin testing; M.tuberculosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..208
FT                   /*tag= a
FT                   /product= Antigen_TbRa32
FT                   /transl_except= (pos:126..127,aa:Pro)
FT                   /note= "Pro coded for by cc; no stop codon given"
XX
PN   WO9709428-A2.
XX
PD   13-MAR-1997.
XX
PF   30-AUG-1996;   96WO-US14674.
XX
PR   12-JUL-1996;   96US-0680574.
PR   01-SEP-1995;   95US-0523436.
PR   22-SEP-1995;   95US-0533634.
PR   22-MAR-1996;   96US-0620874.
PR   05-JUN-1996;   96US-0659683.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-neto A,  Dillon DC,  Houghton R,  Reed SG,  Skeiky YAW;
PI   Twardzik DR, Vedvick TH;
XX
DR   WPI; 1997-192903/17.
DR   P-PSDB; AAW32434.
XX
PT   New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
PT   useful in vaccines for prevention or treatment of tuberculosis, also
PT   for diagnosis
XX
PS   Claim 3; Page 62-63; 168pp; English.
XX
CC   A new immunogenic polypeptide has been developed comprising an
CC   immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
CC   its variant differing only in conservative substitutions and/or
CC   modifications). The present sequence encodes a specifically claimed
CC   M.tuberculosis antigen, TbRa32. The immunogenic protein, and fusion
CC   proteins containing one or more of the proteins or one of the proteins
CC   plus ESAT-6, are useful in vaccines, preferably when formulated with a
CC   non-specific adjuvant, to induce an immune response against
CC   M.tuberculosis (for treatment or prevention).
XX
SQ   Sequence 913 BP; 95 A; 417 C; 278 G; 118 T; 5 other;

  Query Match             1.5%;  Score 49.6;  DB 18;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.0031;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy    2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
            ||| |  |    || |||||||||         |||||  | | |  || |    ||
Db     605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy    2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |      ||  |  |  |   ||        ||||     |||   |  |||
Db     665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy    2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
             |||| |  ||||    || | | |      |:||   | ||       || |     |  ||
Db     725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy    2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
             | || |  |||||||||    |      |   | || || ||     |||    |  || |  |
Db     785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy    2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
            |||| | ||| |  ||||| |   | ||||||||    ||||    ||
Db     845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 32
AAT91413
ID   AAT91413 standard; DNA; 913 BP.
XX
AC   AAT91413;
XX
DT   12-JAN-1998  (first entry)
XX
DE   Mycobacterium tuberculosis antigen TbRa32 encoding DNA.
XX
KW   Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW   skin testing; M.tuberculosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..208
FT                   /*tag= a
FT                   /product= Antigen_TbRa32
FT                   /transl_except= (pos:126..127,aa:Pro)
FT                   /note= "Pro coded for by cc; no stop codon given"
XX
PN   WO9709429-A2.
XX
PD   13-MAR-1997.
XX
PF   30-AUG-1996;   96WO-US14675.
XX
PR   12-JUL-1996;   96US-0680573.
PR   01-SEP-1995;   95US-0523435.
PR   22-SEP-1995;   95US-0532136.
PR   22-MAR-1996;   96US-0620280.
PR   05-JUN-1996;   96US-0658800.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-neto A,  Dillon DC,  Houghton R,  Reed SG,  Skeiky YAW;
PI   Twardzik DR, Vedvick TH;
XX
DR   WPI; 1997-192904/17.
DR   P-PSDB; AAW32366.
XX
PT   New immunogenic polypeptide(s) from soluble M. tuberculosis antigens
PT   - useful for diagnosis of M. tuberculosis infection
XX
PS   Claim 3; Page 64-65; 190pp; English.
XX
CC   A new immunogenic polypeptide has been developed comprising an
CC   immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
```

```
CC   its variant differing only in conservative substitutions and/or
CC   modifications). The present sequence encodes a specifically claimed
CC   M.tuberculosis antigen, TbRa32. The immunogenic polypeptide can be
CC   used to diagnose M.tuberculosis infection by forming complexes with
CC   specific antibodies in the sample. Fragments of DNA encoding the
CC   immunogenic polypeptide can be used as diagnostic primers or probes
CC   and agents that bind to the antigen, especially monoclonal antibodies
CC   or equivalent polyclonal antibodies, are also used for diagnosis.
XX
SQ   Sequence 913 BP; 95 A; 417 C; 278 G; 118 T; 5 other;

  Query Match             1.5%;  Score 49.6;  DB 18;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.0031;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy   2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
           ||| |  |    || |||||||||         |||||  | | |  || |     ||
Db    605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy   2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |         || |  |  |  ||       ||||     |||   |  |||
Db    665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy   2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
             |||| |  ||||   || | | |    |:||    | ||    || |    |   | ||
Db    725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy   2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
            | || |  |||||||||    |      |  | || || ||     |||    |   || |  |
Db    785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy   2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
           |||| | ||| |  ||||| |  | ||||||||   ||||    ||
Db    845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 33
AAV64462
ID   AAV64462 standard; DNA; 913 BP.
XX
AC   AAV64462;
XX
DT   27-JAN-1999  (first entry)
XX
DE   M. tuberculosis immunogenic polypeptide TbRa32 DNA.
XX
KW   Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KW   vaccine; pharmaceutical; infection; diagnosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9816646-A2.
XX
PD   23-APR-1998.
XX
PF   07-OCT-1997;   97WO-US18293.
XX
PR   13-MAR-1997;   97US-0818112.
PR   11-OCT-1996;   96US-0730510.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-Neto A,  Dillon DC,  Houghton R,  Lodes MJ;
PI   Reed SG,  Skeiky YAW,  Twardzik DR,  Vedvick TS;
XX
DR   WPI; 1998-261042/23.
DR   P-PSDB; AAW81669.
XX
PT   Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
PT   to develop products for the detection of M. tuberculosis infection
PT   and for diagnosis, treatment and prevention of tuberculosis
XX
PS   Claim 3; Page 72; 230pp; English.
```

```
XX
CC   This sequence encodes an immunogenic portion of a soluble Mycobacterium
CC   tuberculosis (MT) antigen which can be used in a method for inducing
CC   protective immunity against tuberculosis (TB). This sequence can be
CC   formulated into vaccines and/or pharmaceutical compositions for
CC   immunising against M. tuberculosis infection or may be used for the
CC   diagnosis of tuberculosis.
XX
SQ   Sequence 913 BP; 95 A; 417 C; 278 G; 118 T; 5 other;

  Query Match             1.5%;  Score 49.6;  DB 19;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.0031;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy   2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
           ||| |  |    || |||||||||         |||||  | | |  || |     ||
Db    605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy   2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |         || |  |  |  ||       ||||     |||   |  |||
Db    665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy   2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
             |||| |  ||||   || | | |    |:||    | ||    || |    |   | ||
Db    725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy   2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
            | || |  |||||||||    |      |  | || || ||     |||    |   || |  |
Db    785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy   2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
           |||| | ||| |  ||||| |  | ||||||||   ||||    ||
Db    845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 34
AAV44354
ID   AAV44354 standard; DNA; 913 BP.
XX
AC   AAV44354;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Mycobacterium tuberculosis antigen TbRa32 DNA.
XX
KW   Tuberculosis; infection; diagnosis; antigen; TbRa32; ss.
XX
OS   Mycobacterium tuberculosis strain H37Ra.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..208
FT                   /*tag=  a
FT                   /transl_except= (pos:126..127, aa:Pro)
FT                   /note= "this codon has an apparent 1 nucleotide
FT                          deletion, which alters the reading frame"
XX
PN   WO9816645-A2.
XX
PD   23-APR-1998.
XX
PF   07-OCT-1997;   97WO-US18214.
XX
PR   13-MAR-1997;   97US-0818111.
PR   11-OCT-1996;   96US-0729622.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-Neto A,  Dillon DC,  Houghton R,  Lodes MJ;
PI   Reed SG,  Skeiky YAW,  Twardzik DR,  Vedvick TS;
XX
DR   WPI; 1998-251292/22.
DR   P-PSDB; AAW64306.
XX
```

```
PT   New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT   to develop products for the detection of M. tuberculosis infection
PT   and diagnosis of tuberculosis
XX
PS   Claim 3; Page 71-72; 250pp; English.
XX
CC   This DNA sequence codes for Mycobacterium tuberculosis soluble
CC   antigen TbRa32 (see AAW64306). It was isolated from a M. tuberculosis
CC   strain H37Ra expression library with rabbit anti-sera raised
CC   against M. tuberculosis supernatant.  No significant homology was
CC   found between TbRa32 and Genebank database sequences. The invention
CC   relates to compositions and methods for diagnosing tuberculosis.
CC   It provides polypeptides (see AAW64291-W64379) comprising an
CC   antigenic portion of a soluble M. tuberculosis antigen, or an
CC   immunogenic portion of an M. tuberculosis antigen, as well as DNA
CC   sequences encoding such polypeptides, recombinant expression
CC   vectors and transformed or transfected host cells.  Also claimed
CC   are methods and diagnostic kits for detecting M. tuberculosis
CC   infection in a patient using the above polypeptides, antibodies,
CC   or oligonucleotide probes and primers, for the diagnosis of
CC   tuberculosis.
XX
SQ   Sequence 913 BP; 95 A; 417 C; 278 G; 118 T; 5 other;

  Query Match             1.5%;  Score 49.6;  DB 19;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.0031;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy   2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
          ||| |  |   || |||||||||        |||||  | | |  || |     ||
Db    605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy   2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |         || |  |  |  ||        ||||     |||    |  |||
Db    665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy   2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
            |||| |  ||||   || | | |    |:||    | ||      || |      |  ||
Db    725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy   2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
             | || |  |||||||||    |      |   | || || ||     |||    |  || |  |
Db    785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy   2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
          |||| | ||| |  ||||| |  | ||||||||  ||||   ||
Db    845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 35
AAZ19264
ID   AAZ19264 standard; DNA; 913 BP.
XX
AC   AAZ19264;
XX
DT   05-NOV-1999  (first entry)
XX
DE   M. tuberculosis antigen TbRa32 DNA sequence.
XX
KW   Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW   immunotherapy; diagnosis; immunisation; vaccine; infection;
KW   immune response; skin test; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9942076-A2.
XX
PD   26-AUG-1999.
XX
PF   17-FEB-1999;   99WO-US03268.
XX
PR   05-MAY-1998;   98US-0072967.
PR   18-FEB-1998;   98US-0025197.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-Neto A,  Dillon DC,  Hendrickson RC,  Houghton R;
PI   Lodes MJ,  Reed SG,  Skeiky YAW,  Twardzik DR,  Vedvick TS;
XX
DR   WPI; 1999-527409/44.
DR   P-PSDB; AAY39108.
XX
PT   New antigens from Mycobacterium tuberculosis useful in diagnostic
PT   skin tests and protective or therapeutic vaccines or compositions
XX
PS   Claim 3; Page 79; 299pp; English.
XX
CC   The present invention describes polypeptides comprising an immunogenic
CC   part of a Mycobacterium tuberculosis antigen (Ag). Also described
CC   are vaccines and fusion protein containing M. tuberculosis Ag's.
CC   M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
CC   other polypeptides fragments, can be used in pharmaceutical compositions
CC   or vaccines to generate a protective or therapeutic immune response to
CC   M. tuberculosis and as reagents in skin tests for diagnosis of
CC   tuberculosis. Ag can induce proliferation of, or cytokine secretion
CC   by, T, B or natural killer cells and/or macrophages in
CC   tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAY39083 to
CC   AAY39225 are used in the exemplification of the present invention.
XX
SQ   Sequence 913 BP; 95 A; 417 C; 278 G; 118 T; 5 other;

  Query Match             1.5%;  Score 49.6;  DB 20;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.0031;
  Matches  136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy   2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
          ||| |  |   || |||||||||        |||||  | | |  || |     ||
Db    605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy   2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |         || |  |  |  ||        ||||     |||    |  |||
Db    665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy   2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
            |||| |  ||||   || | | |    |:||    | ||      || |      |  ||
Db    725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy   2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
             | || |  |||||||||    |      |   | || || ||     |||    |  || |  |
Db    785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy   2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
          |||| | ||| |  ||||| |  | ||||||||  ||||   ||
Db    845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 36
AAZ19052
ID   AAZ19052 standard; DNA; 913 BP.
XX
AC   AAZ19052;
XX
DT   05-NOV-1999  (first entry)
XX
DE   M. tuberculosis recombinant antigen DNA encoding TbRa32.
XX
KW   Antigen; diagnosis; detection; infection; antibody; immunisation;
KW   vaccine; immunity; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9942118-A2.
XX
PD   26-AUG-1999.
XX
PF   17-FEB-1999;   99WO-US03265.
```

```
XX
PR   05-MAY-1998;   98US-0072596.
PR   18-FEB-1998;   98US-0024753.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-Neto A,  Dillon DC,  Hendrickson RC,  Houghton R;
PI   Lodes MJ,  Reed SG,  Skeiky YAW,  Twardzik DR,  Vedvick TS;
XX
DR   WPI; 1999-527416/44.
DR   P-PSDB; AAY38971.
XX
PT   New polypeptide comprising antigenic portions of M. tuberculosis
XX
PS   Claim 3 ; Page 112-113; 323pp; English.
XX
CC   This invention describes novel recombinant antigens and their encoding
CC   nucleic acids derived from Mycobacterium tuberculosis. The novel
CC   polypeptides are useful for detecting M. tuberculosis infection in a
CC   biological sample by detecting antibodies which bind with the
CC   polypeptides, and are useful as vaccines for immunizing against
CC   M. tuberculosis infection. The new detection methods are needed as
CC   current vaccination strategies do not provide 100% immunity.
XX
SQ   Sequence 913 BP; 95 A; 417 C; 278 G; 118 T; 5 other;

  Query Match              1.5%;  Score 49.6;  DB 20;  Length 913;
  Best Local Similarity   47.7%;  Pred. No. 0.0031;
  Matches 136;  Conservative    1;  Mismatches  148;  Indels    0;  Gaps    0;

Qy   2065 GCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTG 2124
          ||| |  |    || |||||||||         |||||  | | |  || |     ||
Db    605 GCGCCGGAGNGCGTGCCCGCCGGCGCCGCCAACGCCCAAAAGCCCGGGGTTGCCACCGGC 664

Qy   2125 GAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAA 2184
                |         || |  |  |  ||        ||||     |||   |  |||
Db    665 CCCGCCGGACCCACCGGTCCCGCCGATCCCCCCGTTGCCGCCGGTGCCGCCGCCATTGGT 724

Qy   2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
           |||| |  ||||   || | | |     |:||   | ||     || |    |   | ||
Db    725 GCTGCTGAAGCCGTTAGCGCCGGTTCCGCSGGTTCCGGCGGTGGCGCCNTGGCCGCCGGC 784

Qy   2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
           |  | |  ||||||||||  |     | |  | || ||    |||   |  || |  |
Db    785 CCCGCCGTTGCCGTACAGCCACCCCCCGGTGGCGCCGTTGCCGCCATTGCCGCCATTGCC 844

Qy   2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATG 2349
          |||| | ||| |   ||||| |   | ||||||||   ||||   ||
Db    845 GCCGTTGCCGCCATTGCCGCCGTTCCCGCCGCCACCGCCGGNTTG 889


RESULT 37
AAI99683
ID   AAI99683 standard; DNA; 4403765 BP.
XX
AC   AAI99683;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

  Query Match              1.5%;  Score 49.6;  DB 22;  Length 4403765;
  Best Local Similarity   44.5%;  Pred. No. 0.23;
  Matches 196;  Conservative    0;  Mismatches  244;  Indels    0;  Gaps    0;

Qy      2076 CGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGAT 2135
             |||  |||||| |           |||||  | |   |  ||   || |    ||     | |
Db    336988 CGCTGCCGCCGCTGCCGCCGGCGCCGCCGGAGGCGCCGAGGGCGAGTAGGCCGGCGTTGC 337047

Qy      2136 GGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGC 2195
              | ||      |  |   |||    |    ||  |  | |   | |      || ||    ||
Db    337048 CGCCGCTGCCGCCGTTGCCGCCGAAGGTGCCGCCGCTGCTGCCGCCAGCACCGCCAGTGC 337107

Qy      2196 CGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGC 2255
              || |  | | ||    |||  | |  | |  | ||||  || |||        |  ||       ||
Db    337108 CGCCGGCGCCGAACAGCCCGCCGTGCCCCCCGGCGCCGCCGTCGGCGCCGAGCGTGCCCG 337167

Qy      2256 CGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGT 2315
              |          |||    | ||||||  |       |  |   |    |   | ||| ||||  |   ||
Db    337168 CCCCGCCGGTGCCGCCGGCGCCGAAGAGCAATCCGTTCCCCCCGGTCCCGCCATTCGCGC 337227

Qy      2316 CTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATC 2375
              | |  ||| | |  |||||| ||  ||||  |       |     |  |         |   ||| ||
Db    337228 CAAACCCGCCGGCCCCGCCGGCCCCGCCGTTGGCGAACAGCCCACCGGTACCACCGGCTC 337287

Qy      2376 GGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGG 2435
              | |||     |  |  || || ||  | | || |  |      ||   |     ||             | |
Db    337288 CGGCGGTGCCGCCGGCACCGATAAAGTTTTGGGAGAGGGCGGCCTGGCCGCCGGTCCCTG 337347

Qy      2436 TGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGG 2495
              || ||     |||||| | ||| ||        | |             ||||      ||       |    |
Db    337348 CGGCACCGAGGAACAAGCCGGCGTCACCGCCGCGCCCGCCGGCCCCGCCGGTGTCCAGGC 337407

Qy      2496 CATTGCCGCGAAAACCGGCG 2515
              ||    ||||       ||| ||
Db    337408 CAAACCCGCCGCTGCCGCCG 337427


RESULT 38
AAI99682
ID   AAI99682 standard; DNA; 4411529 BP.
XX
AC   AAI99682;
XX
```

```
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

  Query Match             1.5%;  Score 49.6;  DB 22;  Length 4411529;
  Best Local Similarity   44.1%;  Pred. No. 0.23;
  Matches  208;  Conservative    0;  Mismatches  264;  Indels    0;  Gaps    0;

Qy     2044 GCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTA 2103
              || |  |   | | | | | || |     |   |||  |||||| |       |||||
Db   336867 GCCGACCCCGCCGGTCCCGCCGACATTTAAGGCGCTGCCGCCGCTGCCGCCGGCGCCGCC 336926

Qy     2104 GAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC 2163
              | |   | ||   || |   ||    | |    | ||     |  |   |||    |
Db   336927 GGAGGCGCCGAGGGCGAGTAGGCCGGCGTTGCCGCCGCTGCCGCCGTTGCCGCCGAAGGT 336986

Qy     2164 GCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGC 2223
              ||  |  | |   | |       || ||    |||| | || | | ||    ||| | | | | | |
Db   336987 GCCGCCGCTGCTGCCGCCAGCACCGCCAGTGCCGCCGGCGCCGAACAGCCCGCCGTGCCC 337046

Qy     2224 CCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGT 2283
              |||  || | |||       | ||     ||    |        |||    | |||||| |
Db   337047 CCCGGCGCCGCCGTCGGCGCCGAGCGTGCCCGCCCCGCCGGTGCCGCGGCGCCGAAGAG 337106

Qy     2284 GTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCC 2343
                |  | |    |  | ||| ||||  |   ||  | |  ||| |  |   |||||| ||  |||
Db   337107 CAATCCGTTCCCCCCGGTCCCGCCATTCGCGCCAAACCCGCCGGCCCCGCCGGCCCCGCC 337166

Qy     2344 GATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGT 2403
              | |     |     | |        |   |||  || | |||     | |   || || || |  | |
Db   337167 GTTGGCGAACAGCCCACCGGTACCACCGGCTCCGGCGGTGCCGCCGGCACCGATAAAGTT 337226

Qy     2404 CTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAA 2463
              || | |      ||   |    ||          | |  | || ||   |||||| | ||| ||
Db   337227 TTGGGAGAGGGCGGCCTGGCCGCCGGTCCCTGCGGCACCGAGGAACAAGCCGGCGTCACC 337286

Qy     2464 GGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCG 2515
              | |           ||||   ||      |     | ||    ||||      ||| ||
Db   337287 GCCGCGCCCGCCGGCCCCGCCGGTGTCCAGGCCAAACCCGCCGCTGCCGCCG 337338


RESULT 39
ABK74700
ID   ABK74700 standard; DNA; 750 BP.
XX
AC   ABK74700;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Bacillus licheniformis genomic sequence tag (GST) #1991.
XX
KW   Differential gene expression; genomic sequenced tag; GST;
KW   altered culture condition; environmental stress;
KW   physiological provocation; ds.
XX
OS   Bacillus licheniformis.
XX
PN   WO200229113-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US31437.
XX
PR   06-OCT-2000; 2000US-0680598.
PR   27-MAR-2001; 2001US-279526P.
XX
PA   (NOVO ) NOVOZYMES BIOTECH INC.
PA   (NOVO ) NOVOZYMES AS.
XX
PI   Berka R,  Clausen IG;
XX
DR   WPI; 2002-416684/44.
XX
PT   Monitoring differential expression of several genes in first Bacillus
PT   cell relative to expression of same genes in one or more second
PT   Bacillus cells, by using substrate containing Bacillus genomic
PT   sequenced tag array  -
XX
PS   Claim 4; SEQ ID NO 1991; 200pp; English.
XX
CC   The invention describes a method of monitoring differential expression of
CC   genes in a first Bacillus cell relative to expression of the genes in
CC   other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC   isolated from Bacillus cells to a substrate containing array of Bacillus
CC   genomic sequenced tags (GST), examining the array, and determining
CC   relative gene expression by an observed hybridisation reporter signal of
CC   a spot in the array. The method is useful for measuring the expression of
CC   genes in a first Bacillus cell relative to expression of the same genes
CC   in one or more second Bacillus cells. The method is useful for monitoring
CC   global expression of several genes from a Bacillus cell, discovering new
CC   genes, identifying possible functions of unknown open reading frames and
CC   monitoring gene copy number variation and stability. Monitoring changes
CC   in expression of genes may be used to provide a representation of the way
CC   in which Bacillus cells adapt to changes in culture conditions,
CC   environmental stress or other physiological provocation. Extensive
CC   follow-up characterisation is unnecessary, when one spot on an array
CC   equals one gene or one open reading frame, since sequence information is
CC   available. This sequence represents a genomic sequence tag (GST) used in
CC   the method of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
```

```
SQ   Sequence 750 BP; 228 A; 165 C; 198 G; 159 T; 0 other;

  Query Match             1.5%;  Score 48.4;  DB 24;  Length 750;
  Best Local Similarity   58.2%;  Pred. No. 0.0061;
  Matches   85;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy     1445 CGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCC 1504
             | |||||    | | |||  ||| |    ||   || |||||||| || ||||| ||     |
Db      233 CAAAAGGCGGCTTCCATACAGCCGATCGCTGGCGCAACGACATCTCCGGCACCGGAAAGC 292

Qy     1505 TGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAA 1564
             ||| |||||||||        ||      || | ||| | ||   | || || ||| | |||    |
Db      293 TGACCAAAAAAGGGACAGGCGCTTTGAAGCTGGAAGGGGATAATACATATTCCGGCGGTA 352

Qy     1565 CCATTATCGAAGGCGGTTCGCTGGTG 1590
             |    || ||     ||  | || | |
Db      353 CACGGATTGATCAAGGAACACTTGAG 378


RESULT 40
AAI99683/c
ID   AAI99683 standard; DNA; 4403765 BP.
XX
AC   AAI99683;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

  Query Match             1.4%;  Score 45.2;  DB 22;  Length 4403765;
  Best Local Similarity   57.3%;  Pred. No. 3.9;
  Matches  102;  Conservative    0;  Mismatches   73;  Indels    3;  Gaps    1;

Qy      1225 GGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATAC 1284
              || ||    |||||| ||   ||   ||| | ||||    ||||| ||| ||| |     |||| |
Db   4361104 GGCACCAGCTTTTCGGCGGCCTACGTCAGCGGCGTCGCGGCACTGGTGCGGGCCAAATTC 4361045

Qy      1285 CCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGT 1344
              ||    | |||     | || |   ||     | |   | |   ||      |||| ||    |        |
Db   4361044 CCCGAGTTGACTGCCTACCAAGTGATCAACCGGATCGTGCAGTCGGCGCACAATCCGCCT 4360985

Qy      1345 GCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGA 1402
              ||     |||||||||| |||   | || |    || ||| ||||| |||        ||  |||
Db   4360984 GCG---GGCGTGGACAACAAACTGGGGTACGGGCTGGTGGATCCGGTCGCCGCATTGA 4360930

Search completed: January 27, 2003, 18:45:56
Job time : 7502 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:         January 27, 2003, 19:00:49 ; Search time 105 Seconds
                                             (without alignments)
                                             13708.995 Million cell updates/sec

Title:          US-09-830-433A-7
Perfect score:  3204
Sequence:       1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters:      793544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
               1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
               2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
               3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
               4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
               5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
               6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
               7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
               8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
               9:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
              10:  /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
              11:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
              12:  /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
              13:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
              14:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	270.8	8.5	286	9	US-09-928-457-95	Sequence 95, Appl
	2	48.4	1.5	750	10	US-09-974-300-1991	Sequence 1991, Ap
	3	41	1.3	1866	9	US-09-738-626-1317	Sequence 1317, Ap
c	4	41	1.3	3309400	9	US-09-738-626-1	Sequence 1, Appli
c	5	38.4	1.2	1854	9	US-09-894-844-103	Sequence 103, App
	6	37	1.2	1041	10	US-09-815-242-7885	Sequence 7885, Ap
c	7	37	1.2	6483	10	US-09-880-107-3699	Sequence 3699, Ap
c	8	36.8	1.1	1602	10	US-09-833-790-417	Sequence 417, App
c	9	36.4	1.1	2858	9	US-09-991-496-131	Sequence 131, App
	10	36.4	1.1	4074	10	US-09-815-242-7899	Sequence 7899, Ap
	11	35.8	1.1	3891	9	US-10-108-605-348	Sequence 348, App
	12	35.6	1.1	4689	9	US-09-860-846-34	Sequence 34, Appl
	13	35.6	1.1	4689	10	US-09-861-289-34	Sequence 34, Appl
	14	35.6	1.1	36778	9	US-09-860-846-5	Sequence 5, Appli
	15	35.6	1.1	36778	10	US-09-861-289-5	Sequence 5, Appli
c	16	35.4	1.1	453	10	US-09-960-352-11913	Sequence 11913, A
	17	35.4	1.1	935	10	US-09-770-445-373	Sequence 373, App
	18	35.2	1.1	213	10	US-09-864-761-21194	Sequence 21194, A
c	19	35.2	1.1	422	10	US-09-924-035A-476	Sequence 476, App
	20	35.2	1.1	655	10	US-09-770-149-502	Sequence 502, App
c	21	35.2	1.1	1693	10	US-09-867-550-1669	Sequence 1669, Ap
	22	35	1.1	418	10	US-09-960-352-11256	Sequence 11256, A
	23	35	1.1	1761	10	US-09-969-347-2	Sequence 2, Appli
	24	35	1.1	1761	10	US-09-880-107-3665	Sequence 3665, Ap
	25	35	1.1	1904	10	US-09-925-301-594	Sequence 594, App
	26	35	1.1	3126	10	US-09-815-242-4020	Sequence 4020, Ap
c	27	35	1.1	88421	9	US-09-976-059-1	Sequence 1, Appli
	28	34.6	1.1	1296	10	US-09-815-242-4090	Sequence 4090, Ap
	29	34.4	1.1	3786	10	US-09-815-242-7865	Sequence 7865, Ap
c	30	34.2	1.1	756	9	US-09-938-842A-2617	Sequence 2617, Ap
	31	34.2	1.1	1294	10	US-09-850-964-4	Sequence 4, Appli
	32	34.2	1.1	1488	10	US-09-815-242-7862	Sequence 7862, Ap
	33	34.2	1.1	2100	9	US-09-738-626-1740	Sequence 1740, Ap
	34	34.2	1.1	2289	10	US-09-850-964-3	Sequence 3, Appli
	35	34.2	1.1	3042	9	US-09-712-363-68	Sequence 68, Appl
	36	34.2	1.1	3309400	9	US-09-738-626-1	Sequence 1, Appli
	37	34	1.1	644	10	US-09-867-550-875	Sequence 875, App
c	38	34	1.1	2936	9	US-10-108-605-54	Sequence 54, Appl
c	39	34	1.1	9210	9	US-09-712-363-100	Sequence 100, App
	40	33.8	1.1	969	10	US-09-815-242-7933	Sequence 7933, Ap
	41	33.8	1.1	1768	10	US-09-833-381-523	Sequence 523, App
	42	33.8	1.1	2089	9	US-10-051-902-25	Sequence 25, Appl
	43	33.8	1.1	2089	9	US-10-051-909-25	Sequence 25, Appl
	44	33.8	1.1	2340	10	US-09-815-242-7995	Sequence 7995, Ap
c	45	33.8	1.1	7414	10	US-09-764-853-876	Sequence 876, App

ALIGNMENTS

```
RESULT 1
US-09-928-457-95
; Sequence 95, Application US/09928457
; Patent No. US20020164603A1
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA, specific proteins and peptides
;    TITLE OF INVENTION:  of the Neisseria meningitidis species bacteria, method
;    TITLE OF INVENTION:  for obtaining them and their biological application.
;    NUMBER OF SEQUENCES:  99
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (OEB)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/928,457
;      FILING DATE:  2001-08-14
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/214,759
;      FILING DATE:  199-12-10
;  INFORMATION FOR SEQ ID NO:  95:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  286 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-928-457-95

  Query Match            8.5%;  Score 270.8;  DB 9;  Length 286;
  Best Local Similarity   99.0%;  Pred. No. 3.8e-74;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy        925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy    985 AAAAACATGCTTTTCATTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACAC-ACT 1043
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    121 AAAAACATGCTTTTCATTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACAACT 180

Qy   1044 GACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGT 1103
          |||||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||
Db    181 GACCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTTGCAGGCGT 240

Qy   1104 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 1149
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 286


RESULT 2
US-09-974-300-1991
; Sequence 1991, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1991
;   LENGTH: 750
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-1991

  Query Match             1.5%;  Score 48.4;  DB 10;  Length 750;
  Best Local Similarity   58.2%;  Pred. No. 0.00011;
  Matches   85;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy   1445 CGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCC 1504
           | |||||     | | |||  ||| |    ||    || |||||||| || ||||| ||     |
Db    233 CAAAAGGCGGCTTCCATACAGCCGATCGCTGGCGCAACGACATCTCCGGCACCGGAAAGC 292

Qy   1505 TGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAA 1564
          ||| |||||||||       ||      || | ||| | ||    | || || ||| | |||     |
Db    293 TGACCAAAAAAGGGACAGGCGCTTTGAAGCTGGAAGGGGATAATACATATTCCGGCGGTA 352

Qy   1565 CCATTATCGAAGGCGGTTCGCTGGTG 1590
           |      || ||      ||   | || | |
Db    353 CACGGATTGATCAAGGAACACTTGAG 378


RESULT 3
US-09-738-626-1317
; Sequence 1317, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1317
;   LENGTH: 1866
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1317

  Query Match             1.3%;  Score 41;  DB 9;  Length 1866;
  Best Local Similarity   58.7%;  Pred. No. 0.04;
  Matches   71;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy   1469 ACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAAC 1528
           || ||  |  ||||||||     ||| | ||    ||||| ||| || ||     | |  |||
Db   1340 ACAACTGGCAAAACGACATCGAGGGCGCAGGATCCCTGACCAAGAACGGTTCCGGTGAAC 1399

Qy   1529 TGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGG 1588
           |     |||   ||  ||||| |||| || ||   |||||  |||    |  ||  |  |||
Db   1400 TCACCCTGTCAGGTGACAACTCCTACACCGGTGGAACCACCATCACCGCGGGCACCTTGG 1459

Qy   1589 T 1589
          |
Db   1460 T 1460


RESULT 4
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
;   LENGTH: 3309400
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

  Query Match             1.3%;  Score 41;  DB 9;  Length 3309400;
  Best Local Similarity   58.7%;  Pred. No. 5.6;
  Matches   71;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy   1469 ACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAAC 1528
           || ||  |  ||||||||     ||| | ||    ||||| ||| || ||     | |  |||
Db 1255263 ACAACTGGCAAAACGACATCGAGGGCGCAGGATCCCTGACCAAGAACGGTTCCGGTGAAC 1255204
```

```
Qy     1529 TGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGG 1588
              |    |||    ||   ||||| |||| || ||    |||||  |||     |  ||   |   |||
Db 1255203 TCACCCTGTCAGGTGACAACTCCTACACCGGTGGAACCACCATCACCGCGGGCACCTTGG 1255144

Qy     1589 T 1589
            |
Db 1255143 T 1255143
```

RESULT 5
US-09-894-844-103/c
; Sequence 103, Application US/09894844
; Patent No. US20020176873A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; TITLE OF INVENTION: the M. Tuberculosis Complex
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 1854
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-103

```
  Query Match             1.2%;  Score 38.4;  DB 9;  Length 1854;
  Best Local Similarity   44.1%;  Pred. No. 0.26;
  Matches  208;  Conservative    0;  Mismatches  261;  Indels    3;  Gaps    1;

Qy    2061 GGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAA 2120
             ||    | || |   ||    ||||||| |    |||||    |    |    | | || | ||
Db    1708 GGGCACCGCTCCCGCCAGCCCCGCCGTTGCCAAACAGAACCGCGCCGAAGCCGCCGGCTC 1649

Qy    2121 TCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGT 2180
                 |   | |         | ||       |    |  | | ||        ||||        ||||    |
Db    1648 CGCCGGCACCGATGCCGCCGGCGCCGCCGGCCCCGCCCGCCCCGCCGTTGCCGAACAGCA 1589

Qy    2181 TGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTT 2240
                |  ||| ||   |   |  |  |  |          |       ||||||| |||||| ||
Db    1588 GTGCCGCGTTGCCACCCGCACCGCCGTTGCCGACGCTGCCCGCCCCGCCGGCGCCACCGC 1529

Qy    2241 CCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAG 2300
             | ||| |   ||  ||||       |           |   |||||| ||    |           |   |   |  |
Db    1528 TCCCGCCGGTGCCACCGTTGCCCCACACCAACCCGCCGTCGCCGCCGTTTCCACCGGCGG 1469

Qy    2301 TCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCG 2360
             | |        |  |||  |  | ||  |    | ||||||        ||| ||              ||
Db    1468 CGTAGAGCGTCCCGCCGACGCTGGTGGCACCGCCATTACCGCCCATGCCACCCACACCCA 1409

Qy    2361 GCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCA 2420
             | | |   |||    ||  ||    | | |  |    || ||     | |  ||| |  ||| | ||
Db    1408 ACAGCAGCCCGGCGCGACCGCCATTGCCGCCTGTGCCACCCACCGGGGCGACGTCGGGCG 1349

Qy    2421 AACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAG 2480
                  || || |    ||||   |  | |    |    |    |   |     ||        |  | |
Db    1348 G---CACCACGGTCGTGGTGCCAGCGCCTATACCACCGCCGCTACCACCCATGCCACCCG 1292

Qy    2481 TACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCC 2532
             | || ||     | |||   |||| |      |||| |||||    |      ||| |  |
Db    1291 TGCCAAACAGCTGGGCGATGCCCCCCGCACCGCCGAAACCGCCGGAGCCGGC 1240
```

RESULT 6
US-09-815-242-7885
; Sequence 7885, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7885
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1041)
US-09-815-242-7885

```
  Query Match             1.2%;  Score 37;  DB 10;  Length 1041;
  Best Local Similarity   53.9%;  Pred. No. 0.48;
  Matches   76;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy    2720 GCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTG 2779
             |||||| |||       |   ||| ||          |||| |||||   || ||| ||||        ||
Db     230 GCAGCGGCAAGCAGAGCCTGCTGAGCGCCGGCGAACTGGCTGGGGTCGACGTGTTCTTCC 289

Qy    2780 CCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATG 2839
               |||     |       |          || |  ||| | |||          |      |   |  |||||    |          |  ||
Db     290 TCGCCCTGCACGGCGGCACCGGCGAGGACGGCACCCTGCAGGCGCTGCTGGATGCCGGTG 349

Qy    2840 CATTCGCCGAAAAAGGCAGTG 2860
              |||||| |  |   |||||  |
Db     350 GTTTCGCCTATACCGGCAGCG 370
```

RESULT 7
US-09-880-107-3699/c
; Sequence 3699, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14

```
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3699
;   LENGTH: 6483
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 X14487
US-09-880-107-3699

  Query Match             1.2%;  Score 37;  DB 10;  Length 6483;
  Best Local Similarity   47.6%;  Pred. No. 1.6;
  Matches  146;  Conservative    0;  Mismatches  155;  Indels    6;  Gaps    1;

Qy    2044 GCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTA 2103
              || ||  |  ||||| | || || | |  | |||| ||||  |          |||||
Db    5416 GCCGCCGGAGCTGCTGCCGCCGCCGGAGCTGCCGCCCCCGTAGCCGCCGCCGCCGCCGCC 5357

Qy    2104 GAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC 2163
              | |   | |  || |       |        |  | ||   |||  ||| | | | | |||
Db    5356 GGAACTGCCACCACCGTAGCCGCCGCTGGAACTGCCGCCGTGGCCGCCGCTGGAGCTTCC 5297

Qy    2164 GCAACACCCGAGACGGTTGAAACTGCGGCC------GCCGACCGCACAGATATGCCGGGC 2217
              ||  |     |  | | || | || | |||        |||| |   |  ||   | | |
Db    5296 GCCGCCGTGGCCGCCGCTGGAGCTTCCGCCCCCGTAGCCGCCGCCGGAGCTTCCGCCGCC 5237

Qy    2218 ATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCC 2277
                       ||  |  || | |   |  | || | | || |||||     ||     ||||||
Db    5236 GGAGCTTCCGCCTCCGTAGCCGCCGCCGGAACTGCCGCCGTGGCCGCCGCCGTAGCCGCC 5177

Qy    2278 GACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
              | ||| |    |  |  |      |  ||||||    | |    ||||| | | ||||||||
Db    5176 GCCGGAGCTTCCGCCGCCGGAGCTTCCGCCGCCGTAGCCGCCGCCGAAACTTCCGCCGCC 5117

Qy    2338 CATGCCG 2344
                  |||
Db    5116 GCGTCCG 5110
```

```
RESULT 8
US-09-833-790-417/c
; Sequence 417, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Indirias, Carol Y.
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.512
;  CURRENT APPLICATION NUMBER: US/09/833,790
;  CURRENT FILING DATE:  2001-04-11
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 417
;   LENGTH: 1602
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-790-417

  Query Match             1.1%;  Score 36.8;  DB 10;  Length 1602;
  Best Local Similarity   46.7%;  Pred. No. 0.74;
  Matches  154;  Conservative    0;  Mismatches  172;  Indels    4;  Gaps    1;

Qy    2137 GTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCC 2196
              || | |   || ||| | |   |  |||      ||||   |      |  |  || ||||||
Db    1581 GTTGGAGGAGAGGCCGCTGTGACCCCCGGCTGTCCCGCTGCCCCCGCCAGAGCTGCCGCC 1522

Qy    2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
              | || | | |    || | |  |  ||||  |    ||  ||  |  |||      || |||
Db    1521 GCCCCCGCCGCCCGCCCCGCCGCCGCCCCCGCCACTGCCGCTGCCGCCGCCTGAGCCGCC 1462

Qy    2257 GTACAGCATGCGAATGCCGCCGACGGTGTAC----GCATCTTCAACAGTCTCGCCGCTAC 2312
                   |      || | ||||   | |    |  |      ||| |     |||       ||||| |
Db    1461 TCCCGCGCCGCCACTGCCCGAGCCTCCGCCCCGGTGCACCGCGGACACCGCGGCCGCCGC 1402

Qy    2313 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGT 2372
              || |    |||| |      |||| |||   |||| |  | |   |           ||||      |
Db    1401 CGCCGCAGCCGCCGCCGCCGCTGCCGCCGCCGCTGGGGACAGGTTGGAGCTGCTCTGGGG 1342

Qy    2373 ATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 2432
               ||||     |   ||||  || |  |    |||  ||  ||  |         ||       |      ||
Db    1341 CTCGGCGCTGGGGACACCAGCCCCGGGGCGTGGACGACTCATAGCCGGAGCTGGCGGC 1282

Qy    2433 CGGTGGAACGTGGGAACAGGGCGGTGTTGA 2462
              ||| |      |   ||||    |||| |   ||
Db    1281 CGGGGAGGATTCAGAACCCTGCGGGGAGGA 1252
```

```
RESULT 9
US-09-991-496-131/c
; Sequence 131, Application US/09991496
; Patent No. US20020169285A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Brannon, Mark
;  TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE: 210121.420C9
;  CURRENT APPLICATION NUMBER: US/09/991,496
;  CURRENT FILING DATE:  2001-11-20
;  NUMBER OF SEQ ID NOS: 137
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 131
;   LENGTH: 2858
;   TYPE: DNA
;   ORGANISM: Leishmania major and chagasi
US-09-991-496-131

  Query Match             1.1%;  Score 36.4;  DB 9;  Length 2858;
  Best Local Similarity   51.9%;  Pred. No. 1.4;
  Matches   82;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy    3037 ACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCG 3096
              || || || ||       ||||  | ||||| ||| || | |   ||     |  || | |||
Db    1695 ACGTGCCAGCCGCTAAGTCGCACCACGCCGCGCACACGTCCGCGTGTGTGCGTGTGTGCG 1636

Qy    3097 GATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAA 3156
                ||     | | | ||  ||||              ||  | | | ||  |     | |  |||   |     |
Db    1635 TGTGCGCGCTCCTGCGTCGGCGCACTACTTTTTCCGCTTTTCCTTTTCTACGGCGCAGTA 1576

Qy    3157 CAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTA 3194
              ||| ||   |   |||| || |||   ||   ||| || | |
Db    1575 CAGCACTCCCGCCACGGCAGACTCGTGTGCGGAGACGA 1538
```

```
RESULT 10
US-09-815-242-7899
; Sequence 7899, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
```

; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7899
; LENGTH: 4074
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4074)
US-09-815-242-7899

```
  Query Match             1.1%;  Score 36.4;  DB 10;  Length 4074;
  Best Local Similarity   59.8%;  Pred. No. 1.8;
  Matches   61;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy   1477 CGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTG 1536
Db   1246 CGTACCGAGATCGAAGGTCCGGGCGTCCTGAGCAAGGAAGACATCATCGATGTGCTCAAG 1305

Qy   1537 CACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGC 1578
Db   1306 ACCCTCGTCGACATCCGTAACGGCAAGGGCATCGTCGATGAC 1347
```

RESULT 11
US-10-108-605-348
; Sequence 348, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 348
; LENGTH: 3891
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-108-605-348

```
  Query Match             1.1%;  Score 35.8;  DB 9;  Length 3891;
  Best Local Similarity   49.2%;  Pred. No. 2.7;
  Matches   94;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy   2147 ATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAG 2206
Db   3223 ACGCTGCCACCTCATCCGCACCTGGCGCAGCAGCAACAGCTGCCTCCGCATCCGCACCAG 3282

Qy   2207 ATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATG 2266
Db   3283 TTGCCGCCGCAGTCCCAACCGCTAGCGGGTCCCACCGCCCAGCACCAGCCGTTGCCCCCG 3342

Qy   2267 CGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACA 2326
Db   3343 CACGTAGTGTCGTCGCCCAACGGTGCTCCCTTTGCAGCGCCACCGATTTCGCTCACGGGA 3402

Qy   2327 GTACCGCCGCC 2337
Db   3403 GGACTCCCGCC 3413
```

RESULT 12
US-09-860-846-34
; Sequence 34, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-34

```
  Query Match             1.1%;  Score 35.6;  DB 9;  Length 4689;
  Best Local Similarity   45.4%;  Pred. No. 3.6;
  Matches  128;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy   2077 GCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATG 2136
Db   3406 GCCCACCACGGAGCCGAACACCTCCTCCTCGTCAGCCGCAGCGGCGAACAAGCCCCCGGA 3465

Qy   2137 GTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCC 2196
Db   3466 GCCACCCAACTCACCGCCGAACTCACCGCATCGGGCGCCCGCGTCACCATCGCCGCCTGC 3525

Qy   2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
Db   3526 GACGTCGCCGACCCCCACGCCATGCGCACCCTCCTCGACGCCATCCCCGCCGAGACGCCC 3585

Qy   2257 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTC 2316
Db   3586 CTCACCGCCGTCGTCCACACCGCCGGCGCACCGGGCGGCGATCCGCTGGACGTCACCGGC 3645

Qy   2317 TATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGC 2358
Db   3646 CCGGAGGACATCGCCCGCATCCTGGGCGCGAAGACGAGCGGC 3687
```

RESULT 13

```
US-09-861-289-34
; Sequence 34, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT:  Liu, H.
;  APPLICANT:  Xue, Y.
;  APPLICANT:  Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/861,289
;  CURRENT FILING DATE:  2001-05-18
;  PRIOR APPLICATION NUMBER: 09/105,537
;  PRIOR FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 34
;   LENGTH: 4689
;   TYPE: DNA
;   ORGANISM: Streptomyces venezuelae
US-09-861-289-34

  Query Match             1.1%;  Score 35.6;  DB 10;  Length 4689;
  Best Local Similarity   45.4%;  Pred. No. 3.6;
  Matches  128;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy    2077 GCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATG 2136
            || | |  |||      |||||  | |          || |||||        ||  ||
Db    3406 GCCCACCACGGAGCCGAACACCTCCTCCTCGTCAGCCGCAGCGGCGAACAAGCCCCCGGA 3465

Qy    2137 GTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCC 2196
            | |   |         || |||||   ||||| |   ||        ||     |  || ||| |
Db    3466 GCCACCCAACTCACCGCCGAACTCACCGCATCGGGCGCCCGCGTCACCATCGCCGCCTGC 3525

Qy    2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
            |||  | | ||     |  | ||| ||| ||  |  ||   |  ||| ||| | |  | ||
Db    3526 GACGTCGCCGACCCCCACGCCATGCGCACCCTCCTCGACGCCATCCCCGCCGAGACGCCC 3585

Qy    2257 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTC 2316
             |      |        | ||| ||| | ||      |  | |    || | ||  |||| |
Db    3586 CTCACCGCCGTCGTCCACACCGCCGGCGCACCGGGCGGCGATCCGCTGGACGTCACCGGC 3645

Qy    2317 TATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGC 2358
              |  ||||   ||  |  ||  |  ||  |  |  | |  ||
Db    3646 CCGGAGGACATCGCCCGCATCCTGGGCGCGAAGACGAGCGGC 3687


RESULT 14
US-09-860-846-5
; Sequence 5, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT:  Liu, H.
;  APPLICANT:  Xue, Y.
;  APPLICANT:  Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/860,846
;  CURRENT FILING DATE:  2001-05-18
;  PRIOR APPLICATION NUMBER: 09/105,537
;  PRIOR FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 36778
;   TYPE: DNA
;   ORGANISM: Streptomyces venezuelae
US-09-860-846-5

  Query Match             1.1%;  Score 35.6;  DB 9;  Length 36778;
  Best Local Similarity   45.4%;  Pred. No. 14;
  Matches  128;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy    2077 GCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATG 2136
            || | |  |||      |||||  | |          || |||||        ||  ||
Db   30396 GCCCACCACGGAGCCGAACACCTCCTCCTCGTCAGCCGCAGCGGCGAACAAGCCCCCGGA 30455

Qy    2137 GTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCC 2196
            | |   |         || |||||   ||||| |   ||        ||     |  || ||| |
Db   30456 GCCACCCAACTCACCGCCGAACTCACCGCATCGGGCGCCCGCGTCACCATCGCCGCCTGC 30515

Qy    2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
            |||  | | ||     |  | ||| ||| ||  |  ||   |  ||| ||| | |  | ||
Db   30516 GACGTCGCCGACCCCCACGCCATGCGCACCCTCCTCGACGCCATCCCCGCCGAGACGCCC 30575

Qy    2257 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTC 2316
             |      |        | ||| ||| | ||      |  | |    || | ||  |||| |
Db   30576 CTCACCGCCGTCGTCCACACCGCCGGCGCACCGGGCGGCGATCCGCTGGACGTCACCGGC 30635

Qy    2317 TATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGC 2358
              |  ||||   ||  |  ||  |  ||  |  |  | |  ||
Db   30636 CCGGAGGACATCGCCCGCATCCTGGGCGCGAAGACGAGCGGC 30677


RESULT 15
US-09-861-289-5
; Sequence 5, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT:  Liu, H.
;  APPLICANT:  Xue, Y.
;  APPLICANT:  Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/861,289
;  CURRENT FILING DATE:  2001-05-18
;  PRIOR APPLICATION NUMBER: 09/105,537
;  PRIOR FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 36778
;   TYPE: DNA
;   ORGANISM: Streptomyces venezuelae
US-09-861-289-5

  Query Match             1.1%;  Score 35.6;  DB 10;  Length 36778;
  Best Local Similarity   45.4%;  Pred. No. 14;
  Matches  128;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy    2077 GCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATG 2136
            || | |  |||      |||||  | |          || |||||        ||  ||
Db   30396 GCCCACCACGGAGCCGAACACCTCCTCCTCGTCAGCCGCAGCGGCGAACAAGCCCCCGGA 30455

Qy    2137 GTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCC 2196
            | |   |         || |||||   ||||| |   ||        ||     |  || ||| |
Db   30456 GCCACCCAACTCACCGCCGAACTCACCGCATCGGGCGCCCGCGTCACCATCGCCGCCTGC 30515

Qy    2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
            |||  | | ||     |  | ||| ||| ||  |  ||   |  ||| ||| | |  | ||
Db   30516 GACGTCGCCGACCCCCACGCCATGCGCACCCTCCTCGACGCCATCCCCGCCGAGACGCCC 30575

Qy    2257 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTC 2316
             |      |        | ||| ||| | ||      |  | |    || | ||  |||| |
Db   30576 CTCACCGCCGTCGTCCACACCGCCGGCGCACCGGGCGGCGATCCGCTGGACGTCACCGGC 30635

Qy    2317 TATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGC 2358
              |  ||||   ||  |  ||  |  ||  |  |  | |  ||
Db   30636 CCGGAGGACATCGCCCGCATCCTGGGCGCGAAGACGAGCGGC 30677


RESULT 16
US-09-960-352-11913/c
; Sequence 11913, Application US/09960352
```

```
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11913
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 51-LIB3058-002-Q1-K1-E4
US-09-960-352-11913

  Query Match             1.1%;  Score 35.4;  DB 10;  Length 453;
  Best Local Similarity   48.3%;  Pred. No. 0.87;
  Matches   99;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy    2137 GTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCC 2196
             ||||  |    | |||||  |          | |  |   |||    | | |  | |  || |  |  |
Db     396 GTCGTAGCCGTCGTCTCCGGACTTGTTGTACCCGCCGCCGCCGGTGCCACGGCCGTAGTC 337

Qy    2197 GACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCC 2256
             | |  | | | | |   | | |    ||  | | ||  || || |  |  | |      ||  | || |||
Db     336 GTCGCCGCCGCTGCGGCCGTAGTCGTCTCCGCCGCCGGACTTGTTGTAGCCACCGCCGCC 277

Qy    2257 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTC 2316
             |           ||    | || || ||  |  ||     | |    ||    |||||||  |  |  |
Db     276 GGTGCCACGGCCGTAGTCGTCGCCGCCGGACCTGCCGTACCCATCGCCGCCGCTGCGGCC 217

Qy    2317 TATGCCGACAGTACCGCCGCCCATG 2341
                | || ||   |||||    | ||
Db     216 GTAGTCGTCATCGCCGCCCGACTTG 192


RESULT 17
US-09-770-445-373
; Sequence 373, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
;  APPLICANT:  Gorlach, Jorn
;  APPLICANT:  An, Yong-Qiang
;  APPLICANT:  Hamilton, Carol M.
;  APPLICANT:  Price, Jennifer L.
;  APPLICANT:  Raines, Tracy M.
;  APPLICANT:  Yu, Yang
;  APPLICANT:  Rameaka, Joshua G.
;  APPLICANT:  Page, Amy
;  APPLICANT:  Matthew, Abraham V.
;  APPLICANT:  Ledford, Brooke L.
;  APPLICANT:  Woessner, Jeffrey P.
;  APPLICANT:  Haas, William David
;  APPLICANT:  Garcia, Carlos A.
;  APPLICANT:  Kricker, Maja
;  APPLICANT:  Slader, Ted
;  APPLICANT:  Davis, Keith R.
;  APPLICANT:  Allen, Keith
;  APPLICANT:  Hoffman, Neil
;  APPLICANT:  Hurban, Patrick
;  TITLE OF INVENTION:  Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE:  2023US (PARA-012PRV)
;  CURRENT APPLICATION NUMBER: US/09/770,445
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER:  US 60/178,472
;  PRIOR FILING DATE:  2000-01-27
;  NUMBER OF SEQ ID NOS:  999
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 373
;   LENGTH: 935
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-770-445-373

  Query Match             1.1%;  Score 35.4;  DB 10;  Length 935;
  Best Local Similarity   45.7%;  Pred. No. 1.4;
  Matches  123;  Conservative    0;  Mismatches  146;  Indels    0;  Gaps    0;

Qy    2616 ACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAA 2675
            |  | ||  | |      ||| | ||       |  ||   |    | ||||  | ||       |||
Db     516 ATGTCATCACGATCTCGAAAACGAAAGCGTTCAACACTCCTACTCCAGCAGAGAGACCAA 575

Qy    2676 AAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCAC 2735
            || |  |    |||||     ||||||||  ||         | |   || |  |||| | ||||  |
Db     576 AAGCCGGCACAGCCAAGCCACCGGTGGCGAATTTAAGGATGAGGCAAGCGACGACGGAGC 635

Qy    2736 GCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTT 2795
                || || |        || | |   | |   ||  |      |  ||   |  ||     |||
Db     636 CGAGAAGCTGAGCAATCCAGTAGAGGATACCACGGAGGAGAGTGATGTTACCACCGATGA 695

Qy    2796 GACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGG 2855
              |     |||| |  |          |        || ||  ||        ||||      | ||  |    ||
Db     696 AAGCACCGAAAGTGACGGCAGGGTTAACGTGTCCACCAGAGATGTTGGCACCAACTGAGA 755

Qy    2856 CAGTGCTTTGGGCTGGAGCGGCAACAGCC 2884
            ||| |      |   |   |   ||||   ||||
Db     756 CAGCGACGAAGAGTCCAAAGGCATGAGCC 784


RESULT 18
US-09-864-761-21194
; Sequence 21194, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21194
;   LENGTH: 213
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010906.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
;   OTHER INFORMATION: NT HIT: L42512.1, EVALUE 2.00e-03
;   OTHER INFORMATION: SWISSPROT HIT: P06729, EVALUE 4.40e+00
US-09-864-761-21194

  Query Match             1.1%;  Score 35.2;  DB 10;  Length 213;
  Best Local Similarity   57.1%;  Pred. No. 0.61;
  Matches   64;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy       79 GCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGT 138
              ||  ||  ||  |  |  || || |||  | |||  || ||  ||||  ||   | |
Db       62 GCTCCCTCCAGCTTCCTCCCCGACAATTCAACACCCTCCTCCACCACAGGGAGCTGAGCT 121

Qy      139 ATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACG 190
               |   |  || ||||| |||| || |||| | |||| || ||  ||  || |
Db      122 GTGTTCTTCACCAGCACAGCATCATCAGCCACATCATCACCATCATAGTATG 173


RESULT 19
US-09-924-035A-476/c
; Sequence 476, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
;  APPLICANT: Grlach, Jrn
;  TITLE OF INVENTION: Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE: 2011US
;  CURRENT APPLICATION NUMBER: US/09/924,035A
;  CURRENT FILING DATE:  2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/148,784
;  PRIOR FILING DATE: 1999-08-13
;  NUMBER OF SEQ ID NOS: 900
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 476
;   LENGTH: 422
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(422)
;   OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-476

  Query Match             1.1%;  Score 35.2;  DB 10;  Length 422;
  Best Local Similarity   68.1%;  Pred. No. 0.96;
  Matches   49;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      113 CCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAAT 172
            |||||||||||     | |||  ||| || | |  ||||     | |||||||| || ||
Db      311 CCGACTTCAATGGTTAAGTCACTCGTAACGTCCGAGACAGGCCTGGAACAACAGAGAGAT 252

Qy      173 CAGCAGCAGTAT 184
            ||||| || | |
Db      251 CAGCAACATTTT 240


RESULT 20
US-09-770-149-502
; Sequence 502, Application US/09770149
; Patent No. US20020059663A1
; GENERAL INFORMATION:
;  APPLICANT:  Gorlach, Jorn
;  APPLICANT:  An, Yong-Qiang
;  APPLICANT:  Hamilton, Carol M.
;  APPLICANT:  Price, Jennifer L.
;  APPLICANT:  Raines, Tracy M.
;  APPLICANT:  Yu, Yang
;  APPLICANT:  Rameaka, Joshua G.
;  APPLICANT:  Page, Amy
;  APPLICANT:  Matthew, Abraham V.
;  APPLICANT:  Ledford, Brooke L.
;  APPLICANT:  Woessner, Jeffrey P.
;  APPLICANT:  Haas, William David
;  APPLICANT:  Garcia, Carlos A.
;  APPLICANT:  Kricker, Maja
;  APPLICANT:  Slader, Ted
;  APPLICANT:  Davis, Keith R.
;  APPLICANT:  Allen, Keith
;  APPLICANT:  Hoffman, Neil
;  APPLICANT:  Hurban, Patrick
;  TITLE OF INVENTION:  Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE:  2024 (PARA-013PRV)
;  CURRENT APPLICATION NUMBER: US/09/770,149
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER:  60/178,506
;  PRIOR FILING DATE:  2000-01-27
;  NUMBER OF SEQ ID NOS:  999
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 502
;   LENGTH: 655
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-770-149-502

  Query Match             1.1%;  Score 35.2;  DB 10;  Length 655;
  Best Local Similarity   68.1%;  Pred. No. 1.3;
  Matches   49;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      113 CCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAAT 172
            |||||||||||     | |||  ||| || | |  ||||     | |||||||| || ||
Db      352 CCGACTTCAATGGTTAAGTCACTCGTAACGTCCGAGACAGGCCTGGAACAACAGAGAGAT 411

Qy      173 CAGCAGCAGTAT 184
            ||||| || | |
Db      412 CAGCAACATTTT 423


RESULT 21
US-09-867-550-1669/c
; Sequence 1669, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Mehraban, Fuad,
;  APPLICANT:  Conley, Pamela
;  APPLICANT:  Law, Debbie
;  APPLICANT:  Topper, James
;  TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells a
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 21402-013 (Cura-313)
```

```
;  CURRENT APPLICATION NUMBER: US/09/867,550
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: USSN 60/208,427
;  PRIOR FILING DATE: 2000-05-30
;  NUMBER OF SEQ ID NOS: 2125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1669
;   LENGTH: 1693
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-867-550-1669

  Query Match              1.1%;  Score 35.2;  DB 10;  Length 1693;
  Best Local Similarity   46.4%;  Pred. No. 2.4;
  Matches  115;  Conservative    0;  Mismatches  133;  Indels    0;  Gaps    0;

Qy    2149 GCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGAT 2208
            || ||||  |   | | ||  || |       | |    |      | | |  || |||||||| |
Db     442 GCTTCCGGACGACCGGCGTCAGCTTCCCCAGGCCCAGGCCCAGGCCCCCACCGCACACAA 383

Qy    2209 ATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCG 2268
            |  || |||  | || ||    |   || ||    | || |  ||||      |   |  | |||
Db     382 ACCCCTGGCGCCTGCTCCCCGGAGGCCACGAGCGGCCCCACGGCCTAGGCCGAAAACGCG 323

Qy    2269 AATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGT 2328
                | || || |      |     |   ||       | | | ||| |          ||| |
Db     322 CGGTCGGCGGAGGAGAGGCCTTCCCGCAGGCCGCCCCCGGCTCCGCCTCCCCCGCCGCC 263

Qy    2329 ACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGAC 2388
             |||||||   | |     | || ||   |     | ||      |   |   |||||     ||    |
Db     262 GCCGCCGCCTCGGACACAAAGCCGGCGCTAGGCCTCGGCGAGGCCTCGGAGCACTTCCTC 203

Qy    2389 CACAACGC 2396
             || | |
Db     202 ACCACCCC 195


RESULT 22
US-09-960-352-11256
; Sequence 11256, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11256
;   LENGTH: 418
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 48-LIB34-016-Q1-E1-D8
US-09-960-352-11256

  Query Match              1.1%;  Score 35;  DB 10;  Length 418;
  Best Local Similarity   51.6%;  Pred. No. 1.1;
  Matches   80;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy    2272 GCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACC 2331
            | || | ||||      |||| | |||        |   |||    || | ||||   |  |   ||
Db      39 GTCGACAACGGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCTCCCCGGGCC 98

Qy    2332 GCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCAC 2391
            | |  ||    ||      ||  | | | |||||   | | ||||   ||   ||   ||| ||||
Db      99 GTCTTCCCGTCCATCGTGGGGCGCCCCCGGCACCAGGGCGTAATGGTGGGCATGGGCCAC 158

Qy    2392 AACGCTACGGGTCTGCGCGTCATCGCGCAAACCCA 2426
            || |    ||      || | |        || || | | |
```

```
Db     159 AAGGACTCGTACGTGGGGGATGAGGCTCAGAGCAA 193

RESULT 23
US-09-969-347-2
; Sequence 2, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
;  APPLICANT: Ebner, Reinhard
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-69
;  CURRENT APPLICATION NUMBER: US/09/969,347
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,598
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,604
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 318
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 1761
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-347-2

  Query Match              1.1%;  Score 35;  DB 10;  Length 1761;
  Best Local Similarity   51.6%;  Pred. No. 2.9;
  Matches   80;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy    2272 GCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACC 2331
            | || | ||||      |||| | |||        |   |||    || | |||||   |  |   ||
Db      69 GTCGACAACGGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCCGGGCC 128

Qy    2332 GCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCAC 2391
            | |  ||    ||      ||  | | | | |||    | | | |||    ||   ||    |||   ||
Db     129 GTCTTCCCCTCCATCGTGGGGCGCCCCAGGCACCAGGGCGTGATGGTGGGCATGGGTCAG 188

Qy    2392 AACGCTACGGGTCTGCGCGTCATCGCGCAAACCCA 2426
            || | | |     | || |||  |    || || | | |
Db     189 AAGGATTCCTATGTGGGCGACGAGGCCCAGAGCAA 223


RESULT 24
US-09-880-107-3665
; Sequence 3665, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3665
;   LENGTH: 1761
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 X00351
US-09-880-107-3665

  Query Match              1.1%;  Score 35;  DB 10;  Length 1761;
  Best Local Similarity   51.6%;  Pred. No. 2.9;
  Matches   80;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;
```

```
Qy    2272 GCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACC 2331
           | || |  ||||      |||| | |||        |  |||    || |  |||||  | |  ||
Db      69 GTCGACAACGGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCCGGGCC 128

Qy    2332 GCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCAC 2391
           | |  ||   ||      ||  | | | | | |||    | | |||    ||  ||  |||  ||
Db     129 GTCTTCCCCTCCATCGTGGGGCGCCCCAGGCACCAGGGCGTGATGGTGGGCATGGGTCAG 188

Qy    2392 AACGCTACGGGTCTGCGCGTCATCGCGCAAACCCA 2426
           || | |    | || ||| |    || || | | |
Db     189 AAGGATTCCTATGTGGGCGACGAGGCCCAGAGCAA 223
```

RESULT 25
US-09-925-301-594
; Sequence 594, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 594
;   LENGTH: 1904
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1878)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1893)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1895)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1903)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-594

```
  Query Match             1.1%;  Score 35;  DB 10;  Length 1904;
  Best Local Similarity   51.6%;  Pred. No. 3;
  Matches   80;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy    2272 GCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACC 2331
           | || |  ||||      |||| | |||        |  |||    || |  |||||  | |  ||
Db     142 GTCGACAACGGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCCGGGCC 201

Qy    2332 GCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCAC 2391
           | |  ||   ||      ||  | | | | | |||    | | |||    ||  ||  |||  ||
Db     202 GTCTTCCCCTCCATCGTGGGGCGCCCCAGGCACCAGGGCGTGATGGTGGGCATGGGTCAG 261

Qy    2392 AACGCTACGGGTCTGCGCGTCATCGCGCAAACCCA 2426
           || | |    | || ||| |    || || | | |
Db     262 AAGGATTCCTATGTGGGCGACGAGGCCCAGAGCAA 296
```

RESULT 26
US-09-815-242-4020
; Sequence 4020, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4020
;   LENGTH: 3126
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4020

```
  Query Match             1.1%;  Score 35;  DB 10;  Length 3126;
  Best Local Similarity   61.5%;  Pred. No. 4.2;
  Matches   56;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy    1517 GCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAG 1576
           |||||| |||  |||    ||||    | |||||||| |    |   |       ||| | |||||
Db    1481 GCGGCAACCAGGTGCGGGTGCAGACCGACAACACCGACCAGCCCTTCGACATCAACGAAG 1540

Qy    1577 GCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
           || |   ||||| ||||   |   || |   |||
Db    1541 GCAGCAAGCTGGAGTTGGGCTCCACCGCCAA 1571
```

RESULT 27
US-09-976-059-1/c
; Sequence 1, Application US/09976059
; Patent No. US20020164747A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin
; FILE REFERENCE: 3019-PCT
; CURRENT APPLICATION NUMBER: US/09/976,059
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 88421
;   TYPE: DNA
;   ORGANISM: Actinoplanes sp.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2077)..(3078)
;   OTHER INFORMATION: ORF 1; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (3118)..(4032)
;   OTHER INFORMATION: ORF 2; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (4038)..(5048)

```
;   OTHER INFORMATION: ORF 3; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (4038)..(5048)
;   OTHER INFORMATION: ORF 3; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (6665)..(5814)
;   OTHER INFORMATION: ORF 4; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (7703)..(6693)
;   OTHER INFORMATION: ORF 5; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (9464)..(8130)
;   OTHER INFORMATION: ORF 6; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (9691)..(10761)
;   OTHER INFORMATION: ORF 7; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (12751)..(10829)
;   OTHER INFORMATION: ORF 8; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (13617)..(12802)
;   OTHER INFORMATION: ORF 9; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (15203)..(13614)
;   OTHER INFORMATION: ORF 10; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (15591)..(15863)
;   OTHER INFORMATION: ORF 11; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (15880)..(19035)
;   OTHER INFORMATION: ORF 12; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (19032)..(39713)
;   OTHER INFORMATION: ORF 13; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (39713)..(65800)
;   OTHER INFORMATION: ORF 14; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (65826)..(66530)
;   OTHER INFORMATION: ORF 15; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (66546)..(67370)
;   OTHER INFORMATION: ORF 16; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (67384)..(70059)
;   OTHER INFORMATION: ORF 17; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (70099)..(70662)
;   OTHER INFORMATION: ORF 18; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (70659)..(71906)
;   OTHER INFORMATION: ORF 19; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (73439)..(71964)
;   OTHER INFORMATION: ORF 20; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (74216)..(73563)
;   OTHER INFORMATION: ORF 21; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (75424)..(74213)
;   OTHER INFORMATION: ORF 22; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (75535)..(76464)
;   OTHER INFORMATION: ORF 23; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (78110)..(76449)
;   OTHER INFORMATION: ORF 24; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (79864)..(78107)
;   OTHER INFORMATION: ORF 25; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (81624)..(79861)
;   OTHER INFORMATION: ORF 26; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (81909)..(81682)
;   OTHER INFORMATION: ORF 27; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (82346)..(82062)
;   OTHER INFORMATION: ORF 28; negative strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (82587)..(84446)
;   OTHER INFORMATION: ORF 29; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (84481)..(85548)
;   OTHER INFORMATION: ORF 30; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (85556)..(86845)
;   OTHER INFORMATION: ORF 31; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (87372)..(86803)
;   OTHER INFORMATION: ORF 32; positive strandedness
;   NAME/KEY: misc_feature
;   LOCATION: (87494)..(88420)
;   OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only
US-09-976-059-1

  Query Match             1.1%;  Score 35;  DB 9;  Length 88421;
  Best Local Similarity   45.7%;  Pred. No. 38;
  Matches  122;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy    2239 TTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAAC 2298
           | ||| || ||||  |  ||   |||   | | | | | |  |  |  |    | |  ||
Db   25166 TCCCGTGCCGCGGGCGAGGTCGGGCACCGGCAGGGCCCGGCGGTCGAGCTTGCCGTTCAC 25107

Qy    2299 AGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGC 2358
            |||  |      | ||| |   | ||   | |      ||    |  |  | |  ||
Db   25106 CGTCAAGGGCAACCGGTCCAGCTCCACGAAAGCCGACGGCACCATGTAACCCGGCAACCG 25047

Qy    2359 CGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCG 2418
           | ||||   |    | ||||  | || |||  || |||    |   | |||  |   |||
Db   25046 CTGCTGCGCGTACGACCGGAGAGCTTCGACGTCACCGCCCACGACGTACGCCACCAAGCG 24987

Qy    2419 CAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGC 2478
           |    |   | | |||     |||   || ||  ||||  | |    ||    |  ||||
Db   24986 CTTGTCACCCGGCACGTCCTCACGCACGATCACCGCGGCCTGGCTCACACCCGGTCCGGC 24927

Qy    2479 AGTACCCAAACCGTCGGCATTGCCGCG 2505
            |    ||   | ||  | | |||| |
Db   24926 GGCCAGCACGGCCTCCACCTCGCCGGG 24900


RESULT 28
US-09-815-242-4090
; Sequence 4090, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
```

```
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4090
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4090

  Query Match             1.1%;  Score 34.6;  DB 10;  Length 1296;
  Best Local Similarity   48.7%;  Pred. No. 3.1;
  Matches   94;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy   2626 GCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATC 2685
Db    346 GCGGTCGGCATCGCCTACACCTCGGCCTGGTTCGAGAAGGAGCGCCAGGGCACCGCGATG 405

Qy   2686 AGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAG 2745
Db    406 GGCATCTTCGGCGCCGGCAACGCCGGCGCGGCGATCACCAACCTGGTGGCGCCGATGATC 465

Qy   2746 CTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAA 2805
Db    466 GTGGTCGCCTTCGGCTGGCGCATGGTGCCGCAGGTCTACTCGGTGGCCATGCTGGTCACC 525

Qy   2806 GGCGGTCTGCGCT 2818
Db    526 GCCGTGCTGTTCT 538


RESULT 29
US-09-815-242-7865
; Sequence 7865, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7865
;   LENGTH: 3786
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3786)
US-09-815-242-7865

  Query Match             1.1%;  Score 34.4;  DB 10;  Length 3786;
  Best Local Similarity   54.8%;  Pred. No. 7.3;
  Matches   68;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy     44 CTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCA 103
Db   1713 CGGCCAAAGCGGCGGCGGCTGGGCGCACTATGTCGGCCAGGAGAAGCTGCGCCCGCAGAC 1772

Qy    104 CTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAA 163
Db   1773 CGGCTGGGCGCCGCTGGCCTTCGGCCTGGACTGGAGCCGGCCGCCGCGGCAGATGAACGG 1832

Qy    164 CAGC 167
Db   1833 CACC 1836


RESULT 30
US-09-938-842A-2617/c
; Sequence 2617, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2617
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2617

  Query Match             1.1%;  Score 34.2;  DB 9;  Length 756;
  Best Local Similarity   45.4%;  Pred. No. 2.9;
  Matches  123;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy   2616 ACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAA 2675
Db    457 ATGTCATCACGATCTCGAAAACGAAAGCGTTCAACACTCCTACTCCAGCAGAGAGACCAA 398

Qy   2676 AAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCAC 2735
Db    397 AAGCCGGCACAGCCAAGCCACCGGTGGCGAATTTAAGGATGAGGCAAGCGACGACGGAGC 338

Qy   2736 GCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTT 2795
Db    337 CGAGAAGCTGAGCAATCCAGTAGAGGATACCACGGAGGAGAGTGATGTTACCACCAATGA 278

Qy   2796 GACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGG 2855
Db    277 AAGCACCGAAAGTGACGGCAGGGTTAACGTGTCCACCAGAGATGTTGGCACCAACTGAGA 218

Qy   2856 CAGTGCTTTGGGCTGGAGCGGCAACAGCCTC 2886
Db    217 CAGCGACGAAGAGTCCAAAGGCATGAGCCAC 187
```

```
RESULT 31
US-09-850-964-4
; Sequence 4, Application US/09850964
; Patent No. US20020104121A1
; GENERAL INFORMATION:
;  APPLICANT: MCELROY, DAVID
;  APPLICANT:  KRIZ, ALAN L.
;  APPLICANT:  OROZCO, EMIL
;  APPLICANT:  GRIFFOR, MATT
;  TITLE OF INVENTION: THE MAIZE A3 PROMOTER AND RICE ACTIN 2 INTRON AND
;  TITLE OF INVENTION:  METHODS FOR USE THEREOF
;  FILE REFERENCE: DEKM:166
;  CURRENT APPLICATION NUMBER: US/09/850,964
;  CURRENT FILING DATE:  2001-05-07
;  PRIOR APPLICATION NUMBER: 09/312,038
;  PRIOR FILING DATE: 1999-05-14
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 1294
;   TYPE: DNA
;   ORGANISM: Zea mays/rice
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (883)..(972)
;   OTHER INFORMATION: N = A or C or G or T
US-09-850-964-4

  Query Match             1.1%;  Score 34.2;  DB 10;  Length 1294;
  Best Local Similarity   47.1%;  Pred. No. 4.1;
  Matches  105;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy   2115 CAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGA 2174
          || ||   | |  |||  |||          | |    |    || |     || |    | || |
Db    466 CATCACGGTAGCGAACAGGATAACTTGGCCGCGCGGGCCCCACCGCTCAGTTCCTCCAGT 525

Qy   2175 GACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGC 2234
            | || | |  ||| |  |||| |  ||| | |       | | | |    | ||  |  | |
Db    526 TCCCGTCGGAGTTGCCGTAGCCGCCGCCACGGTTGCCACCGCCGTAGCCGCCACCACCAC 585

Qy   2235 AACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTT 2294
           ||  |    | | || ||  |||||           ||      |||||| |   |     |  | |
Db    586 CACCGTAGCCACCGCCGCCACCGTAGCCGCCGCCGCCACCGCCGTAGCCGCCGCCGCCAT 645

Qy   2295 CAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
          ||      | |||| |  ||||      |||| |    || ||| ||
Db    646 CACGGCGCCCGCCACCGCCGTAGCCGCCGCCTCCACGGCCACC 688


RESULT 32
US-09-815-242-7862
; Sequence 7862, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7862
;   LENGTH: 1488
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1488)
US-09-815-242-7862

  Query Match             1.1%;  Score 34.2;  DB 10;  Length 1488;
  Best Local Similarity   56.8%;  Pred. No. 4.5;
  Matches   63;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy   2652 CCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACA 2711
          |||| |    |  || | || |        ||    || ||  |||| ||    |||| |||  |
Db     51 CCTGGTGCTCGCCGTAGGCGAAGGCCGCAAGCTCGGCGCCAGCGCCAAGGCGGTCGACGA 110

Qy   2712 TGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGG 2762
           ||     ||| ||| || | || |  || |  |||  ||||  |||| |||
Db    111 CGCCACCGGCGGCGCCATCAGCGCAGTGCTCAAGCGTGGCGACCTGGCCGG 161


RESULT 33
US-09-738-626-1740
; Sequence 1740, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1740
;   LENGTH: 2100
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1740

  Query Match             1.1%;  Score 34.2;  DB 9;  Length 2100;
  Best Local Similarity   46.8%;  Pred. No. 5.7;
  Matches  108;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy   2216 GCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCG 2275
          | | ||  | || |||       |   | |||  || | || | |   | |   | | | ||
```

```
Db   1307 GTAACCTGCACTTCGGTATCCGTGAGCACGCTATGGGATCCATCCTCAACGGCATTTCCC 1366

Qy   2276 CCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCG 2335
Db   1367 TCCACGGTGGCACCCGCCCATACGGCGGAACCTTCCTCATCTTCTCCGACTACATGCGTC 1426

Qy   2336 CCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACG 2395
Db   1427 CTGCAGTTCGTCTTGCAGCTCTCATGGAGACCGACGCTTACTACGTCTGGACCCACGACT 1486

Qy   2396 CTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGG 2446
Db   1487 CCATCGGTCTGGGCGAAGATGGCCCAACCCACCAGCCTGTTGAAACCTTGG 1537
```

RESULT 34
US-09-850-964-3
; Sequence 3, Application US/09850964
; Patent No. US20020104121A1
; GENERAL INFORMATION:
; APPLICANT: MCELROY, DAVID
; APPLICANT: KRIZ, ALAN L.
; APPLICANT: OROZCO, EMIL
; APPLICANT: GRIFFOR, MATT
; TITLE OF INVENTION: THE MAIZE A3 PROMOTER AND RICE ACTIN 2 INTRON AND
; TITLE OF INVENTION: METHODS FOR USE THEREOF
; FILE REFERENCE: DEKM:166
; CURRENT APPLICATION NUMBER: US/09/850,964
; CURRENT FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/312,038
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2289
;   TYPE: DNA
;   ORGANISM: Zea mays/rice
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (889)..(1825)
;   OTHER INFORMATION: N = A or C or G or T
US-09-850-964-3

```
  Query Match             1.1%;  Score 34.2;  DB 10;  Length 2289;
  Best Local Similarity   47.1%;  Pred. No. 6;
  Matches  105;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy   2115 CAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGA 2174
Db    472 CATCACGGTAGCGAACAGGATAACTTGGCCGCGCGGGCCCCACCGCTCAGTTCCTCCAGT 531

Qy   2175 GACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGC 2234
Db    532 TCCCGTCGGAGTTGCCGTAGCCGCCGCCACGGTTGCCACCGCCGTAGCCGCCACCACCAC 591

Qy   2235 AACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTT 2294
Db    592 CACCGTAGCCACCGCCGCCACCGTAGCCGCCGCCGCCACCGCCGTAGCCGCCGCCGCCAT 651

Qy   2295 CAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
Db    652 CACGGCGCCCGCCACCGCCGTAGCCGCCGCCTCCACGGCCACC 694
```

RESULT 35
US-09-712-363-68
; Sequence 68, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
;   LENGTH: 3042
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-68

```
  Query Match             1.1%;  Score 34.2;  DB 9;  Length 3042;
  Best Local Similarity   50.9%;  Pred. No. 7.3;
  Matches   81;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy   1231 TCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGG 1290
Db   1102 TCGGGTACGGCCGCGGTGCTCGACGAGGCTGGGTTTGGCGTGCTGCTGCCGTCCTGGTGG 1161

Qy   1291 ATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTC 1350
Db   1162 GACCGCCGCCGCAAGCTGGGCTTGGTCCTGTCCGCATATACCCCGGTCGACGGCGTGGTG 1221

Qy   1351 GGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCG 1389
Db   1222 GGCAAGGCCAGCAAGTTCGGCCGCGAGCAGCTCGTCGAG 1260
```

RESULT 36
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059

```
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
;   LENGTH: 3309400
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

  Query Match             1.1%;  Score 34.2;  DB 9;  Length 3309400;
  Best Local Similarity   46.8%;  Pred. No. 1.3e+02;
  Matches  108;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy    2216 GCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCG 2275
                | | ||  | || |||       |   | |||  || | || | |    | |  | | ||
Db 1665709 GTAACCTGCACTTCGGTATCCGTGAGCACGCTATGGGATCCATCCTCAACGGCATTTCCC 1665768

Qy    2276 CCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCG 2335
             | ||||||    |  |      || |      ||      | |||   |||||    |
Db 1665769 TCCACGGTGGCACCCGCCCATACGGCGGAACCTTCCTCATCTTCTCCGACTACATGCGTC 1665828

Qy    2336 CCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACG 2395
                |    |     | |    | | |  |  ||  | ||  |    |||  | |  |||| ||
Db 1665829 CTGCAGTTCGTCTTGCAGCTCTCATGGAGACCGACGCTTACTACGTCTGGACCCACGACT 1665888

Qy    2396 CTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGG 2446
            | |  |||||| |||      |  | |||||| |||    | || ||| | ||
Db 1665889 CCATCGGTCTGGGCGAAGATGGCCCAACCCACCAGCCTGTTGAAACCTTGG 1665939


RESULT 37
US-09-867-550-875
; Sequence 875, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Mehraban, Fuad,
;  APPLICANT:  Conley, Pamela
;  APPLICANT:  Law, Debbie
;  APPLICANT:  Topper, James
;  TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 21402-013 (Cura-313)
;  CURRENT APPLICATION NUMBER: US/09/867,550
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: USSN 60/208,427
;  PRIOR FILING DATE: 2000-05-30
;  NUMBER OF SEQ ID NOS: 2125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 875
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-867-550-875

  Query Match             1.1%;  Score 34;  DB 10;  Length 644;
  Best Local Similarity   54.2%;  Pred. No. 3;
  Matches   90;  Conservative    0;  Mismatches   75;  Indels    1;  Gaps    1;

Qy    2093 AACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCT 2152
                || | |  | |  | |||  ||| ||||  | ||  ||  | |||  |       |   ||
Db     252 AAAAGGAAGCAAGAGAGGCAGGCTGCAAGTTTGATCACTGGGTGGCTGCTAGCCACCTCT 311

Qy    2153 CCGAATCATCCGCAACACCCGAGAC-GGTTGAAACTGCGGCCGCCGACCGCACAGATATG 2211
                ||    ||| |  ||| | | || || || |||||  || || || | | ||       |
Db     312 CCCCCTCAGCTCCAAGAGCAGAAACAGGATGAAAACGCAGCAGCTGCCGGCTTCTGTCCT 371

Qy    2212 CCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCG 2257
                |  | || | ||| ||   ||   |||   |  | || ||||||||
Db     372 CTTGCCAGCTGCCTCTCTTGCAAGACTAGACAGGGAGGGGCAGCCG 417


RESULT 38
US-10-108-605-54/c
; Sequence 54, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
;  APPLICANT: Broadus, Julie
;  APPLICANT:  Stam, Lynn
;  APPLICANT:  Bachmann, Jane
;  APPLICANT:  Kamdar, Kim
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOD
;  TITLE OF INVENTION:  PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
;  FILE REFERENCE: 31133B
;  CURRENT APPLICATION NUMBER: US/10/108,605
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: US 09/761,142
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/176,418
;  PRIOR FILING DATE: 2000-01-14
;  NUMBER OF SEQ ID NOS: 361
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
;   LENGTH: 2936
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-10-108-605-54

  Query Match             1.1%;  Score 34;  DB 9;  Length 2936;
  Best Local Similarity   48.9%;  Pred. No. 8.2;
  Matches   91;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;

Qy    2151 CTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATAT 2210
            | |||||||  || | ||  || |  |||||    |  | || | |   |     |   | |
Db     699 CCCCGAATCCACCACCACCTCCAAATCGGTTGCCGCCTCCGCGGTCATCGCGTCGCTTCT 640

Qy    2211 GCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAA 2270
                  |||  |   || || | ||        ||  | |  ||  | ||||       | ||||
Db     639 CGACGGCTCCATTCCTTATGCCGTGGTAATCGCCACCGCCACCGCCGCCACCGAAGCGAT 580

Qy    2271 TGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTAC 2330
            |||| ||| ||   |      |          | ||     |||| | |  || ||  ||    |   ||
Db     579 TGCCTCCGCCGCCGCCTCCTCGCCGATCATCGCCGCCTCGATCGCCTCCGCGACCGCCAC 520

Qy    2331 CGCCGC 2336
              || |
Db     519 GTCCAC 514


RESULT 39
US-09-712-363-100/c
; Sequence 100, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE: 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
```

```
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100
;   LENGTH: 9210
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-100

  Query Match              1.1%;  Score 34;  DB 9;  Length 9210;
  Best Local Similarity   44.7%;  Pred. No. 17;
  Matches  176;  Conservative    0;  Mismatches  215;  Indels    3;  Gaps    1;

Qy     2704 GACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGT 2763
Db     2985 GACGGGCACGCTCAAGGCGACGTCGTCGCCGTGCGTCTGCAGCCGGGCGCCGGTGGATGA 2926

Qy     2764 GTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGAC 2823
Db     2925 GTGTGTGGCGCGCGGGTTTTCGGGTCCATCGTGCACCTGCCATTCGGCCGGGTCGGCGAT 2866

Qy     2824 CTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGC 2883
Db     2865 CCGATGCACCGGGTTGGTCACGGTGCGACCGGCCCAGCGCACATCGGGTGCGTCGAGGAC 2806

Qy     2884 CTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGAT 2943
Db     2805 GACAGCCAACGGTCCGGCCACGTCGGCGCGGCCC---AGCCGGCGCGACGCGACATCCTT 2749

Qy     2944 AAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACG 3003
Db     2748 CGGCTCGACACCGGCGCCGAGCACTTCATCGATTGCGGCTTGCTCGAAACGGTCCAGCAA 2689

Qy     3004 GTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATG 3063
Db     2688 CTCACCGACGGGTTCATCCATCCGGGTGATGCCGGCTACCGACGCGGTGCCCGGAATGAT 2629

Qy     3064 CCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGG 3097
Db     2628 GCACACCGCATCGGCGTCGTAGCGGGCGTCGTGG 2595

RESULT 40
US-09-815-242-7933
; Sequence 7933, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7933
;   LENGTH: 969
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(969)
US-09-815-242-7933

  Query Match              1.1%;  Score 33.8;  DB 10;  Length 969;
  Best Local Similarity   47.8%;  Pred. No. 4.5;
  Matches   98;  Conservative    0;  Mismatches  107;  Indels    0;  Gaps    0;

Qy     2725 GTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCA 2784
Db      700 GGCAAGCCCACCCTGCCGCTGATCGTCACCATGCGCGACGGCACCGAGGAACAGGCCGCG 759

Qy     2785 ACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTC 2844
Db      760 CTGGTACGCAAAGCCATCCAGCAAGGTGGCAGCCAGGACCTGGAGAGCGTCTGTGCCGCC 819

Qy     2845 GCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTC 2904
Db      820 GTCGAGGCAGCCGGCGCCCTGGACTATACCGCCAACCTCGCCCGCGACTACGCCGCCCGC 879

Qy     2905 GGACTCGCGGGTCTGAAGCTGTCGC 2929
Db      880 GCCATCGCCTGCCTGGATACGCTGC 904

Search completed: January 28, 2003, 00:26:14
Job time : 14101 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:         January 27, 2003, 16:40:30 ; Search time 2876 Seconds
                                             (without alignments)
                                             18042.538 Million cell updates/sec

Title:          US-09-830-433A-7
Perfect score:  3204
Sequence:       1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	57.4	1.8	910	17	CNS006ON	AL065629 Drosophil
c	2	47.4	1.5	580	14	BQ620456	BQ620456 TaLr1156F
c	3	46.4	1.4	925	17	CNS0091P	AL053013 Drosophil
c	4	45.8	1.4	902	17	CNS006QP	AL065804 Drosophil
	5	45.6	1.4	471	10	AW707761	AW707761 832012E07
	6	44.8	1.4	1101	17	CNS012S8	AL101954 Drosophil
	7	44.6	1.4	584	9	AL824948	AL824948 AL824948
	8	44	1.4	488	12	BE726719	BE726719 894093B05
	9	44	1.4	490	10	AV639324	AV639324 AV639324
	10	44	1.4	497	10	AV636203	AV636203 AV636203
	11	44	1.4	513	10	AV636123	AV636123 AV636123
	12	44	1.4	557	14	BQ822928	BQ822928 1030104H1
	13	44	1.4	562	10	AV387616	AV387616 AV387616
	14	44	1.4	634	13	BI875064	BI875064 963121E11
	15	43.8	1.4	1087	17	CNS0106Y	AL098596 Drosophil
	16	43.8	1.4	1101	17	CNS017X5	AL108611 Drosophil
	17	43.2	1.3	655	10	AW676029	AW676029 832001E09
	18	43	1.3	477	10	AV633578	AV633578 AV633578
	19	42.2	1.3	358	13	BG947100	BG947100 IP1_2_E12
c	20	42.2	1.3	755	10	BE041076	BE041076 OF19D09 O
c	21	42	1.3	1101	17	CNS017GS	AL108022 Drosophil
	22	42	1.3	1101	17	CNS017SY	AL108460 Drosophil
	23	41.8	1.3	536	9	AL716352	AL716352 AL716352
c	24	41.8	1.3	538	14	BQ283304	BQ283304 WHE3088_D
	25	41.8	1.3	650	9	AL721811	AL721811 AL721811
	26	41.8	1.3	766	9	AL721550	AL721550 AL721550
	27	41.8	1.3	772	9	AL716995	AL716995 AL716995
	28	41.8	1.3	790	9	AL716744	AL716744 AL716744
c	29	41.8	1.3	792	9	AL717085	AL717085 AL717085
	30	41.8	1.3	844	17	CNS0052P	AL056652 Drosophil
c	31	41.6	1.3	645	17	CNS012I3	AL101589 Drosophil
	32	41.4	1.3	444	14	BQ759294	BQ759294 EBpi01_SQ
	33	41.4	1.3	976	12	BE778605	BE778605 601466144
	34	41.2	1.3	588	13	BI778455	BI778455 EBro07_SQ
	35	41.2	1.3	847	17	CNS021SP	AL177298 Tetraodon
	36	41	1.3	573	12	BG368419	BG368419 HVSMEi001
	37	40.8	1.3	540	10	AW006888	AW006888 ws15f06.x
c	38	40.8	1.3	631	13	BI790356	BI790356 id06e08.x
	39	40.4	1.3	474	13	BM068591	BM068591 WHE3459_D
	40	40.4	1.3	1600	17	AG058272	AG058272 Pan trogl
c	41	40.2	1.3	497	12	BF657250	BF657250 OV2_22_E0
	42	40.2	1.3	595	9	AL716267	AL716267 AL716267
	43	40.2	1.3	692	17	CNS007WH	AL050923 Drosophil
	44	40.2	1.3	715	9	AL715439	AL715439 AL715439
	45	40.2	1.3	722	9	AL728402	AL728402 AL728402

ALIGNMENTS

```
RESULT 1
CNS006ON/c
LOCUS       CNS006ON                 910 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL065629
VERSION     AL065629.1  GI:4944698
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 910)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
```

```
           EcoRI digestion of Drosophila DNA provided by the BDGP from the
           isogenic strain y2; cn bw sp, the same strain used for the BDGP's
           P1 and EST libraries. A more detailed description of the library
           and how to order individual BAC clones, the entire library, or
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..910
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR14J21"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
BASE COUNT      202 a     63 c    112 g    198 t    335 others
ORIGIN

  Query Match             1.8%;  Score 57.4;  DB 17;  Length 910;
  Best Local Similarity   16.5%;  Pred. No. 6.7e-05;
  Matches   59;  Conservative  159;  Mismatches  139;  Indels    0;  Gaps    0;

Qy   1982 GCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCA 2041
           ::: |  |: ||:||  :  :|: :|:: || : ::: : :: |     :  :: : :::::
Db    869 RSSVCASAMGAASAGASAASASSSAVSACGSSSSSVSSVVSSASASASAGVSRSASSVSV 810

Qy   2042 ATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCG 2101
            : : :::::|: ::   : |:| :: :|:|  ::: : : ::::  :   |:|:| :  :
Db    809 SSVGSASRSRSASVSSSSSAGSAAVSAVAMACASMVASAVASVMSASVVAARASAKSAAS 750

Qy   2102 TAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCAT 2161
            |::| :: :: : |::|  :  :|:||:: :: :   ::|: |:  ::: :::|  :|
Db    749 MARVAAVVARVASAAVVASASAAVAVAAVSRSVASASRVMASAGSSASSSSSSSASSSAS 690

Qy   2162 CCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCC 2221
           :: : |::::: ::|:::   | |  : :: :: : :: : : |:  |   ::|:::: |:
Db    689 SSASSASMVVSCRSAVSVASAASAAASARSASRASAVMAVAVAASSSAVASSGSSSVSCS 630

Qy   2222 GCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACG 2281
           :: |      :::: |:   :::::: :: || :| : |: ::: |:     :::||:: :
Db    629 SSRCGCSASSSSASASSGSSSSSSSSSSCGCGSCASCASSSSRSGSCGCGSSVGCVSSSC 570

Qy   2282 GTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCC 2338
           | :  :::  :      : : | |:|:      |   |    :|:  |     :: |::|||
Db    569 GSSSSSSSSSSSGCSSSSCCCCSCSCACCRCCCCCRSCSCGCCCMSSSCCSSCCC 513


RESULT 2
BQ620456/c
LOCUS       BQ620456                 580 bp    mRNA    linear   EST 28-JUN-2002
DEFINITION  TaLr1156F12R TaLr1 Triticum aestivum cDNA clone TaLr1156F12R, mRNA
            sequence.
ACCESSION   BQ620456
VERSION     BQ620456.1  GI:21625535
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Cloutier,S., Dong,G. and Walsh,A.
  TITLE     Wheat functional genomics - Thatcher Lr1 cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Sylvie Cloutier
            Cereal Research Centre, Agriculture and Agri-food Canada
            195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
            Tel: (204) 983-2340
            Fax: (204) 983-4604
            Email: scloutier@em.agr.ca
            was cloned directionally, not all sequences generated with reverse
            primer where from the 5' end (same with forward primer and 3'end).
            Average insert size is >2.2kb
            Plate: 156  row: F  column: 12
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..580
                     /organism="Triticum aestivum"
                     /cultivar="Thatcher Lr1"
                     /db_xref="taxon:4565"
                     /clone="TaLr1156F12R"
                     /clone_lib="TaLr1"
                     /tissue_type="Leaf tissue"
                     /dev_stage="14 Days old"
                     /lab_host="E. coli XLOLR"
                     /note="Vector: Lambda ZapII; mass excised in plasmid
                     vector pBK-CMV (Stratagene).; Site_1: EcoRI; Site_2: XhoI;
                     mRNA obtained from wheat NIL Thatcher Lr1 24 hours after
                     inoculation  with leaf rust pathogen Puccinia triticina
                     race BBB carrying the  avirulence gene Avr1."
BASE COUNT       77 a    224 c    195 g     84 t
ORIGIN

  Query Match             1.5%;  Score 47.4;  DB 14;  Length 580;
  Best Local Similarity   43.7%;  Pred. No. 0.042;
  Matches  210;  Conservative    0;  Mismatches  271;  Indels    0;  Gaps    0;

Qy   2064 AGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCT 2123
           ||  |  |   |   | ||| |    | ||||  |||     ||||  |      ||| ||   | ||
Db    579 AGTAGGAGCTCACCCGCGCCAGGCTCTGGCACAGCGTGTAGTCCTCCTCGGAAGTCAGCT 520

Qy   2124 GGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGA 2183
           || |  | ||    ||    |        |    ||||   |   |   |    |   |    ||
Db    519 GGTAGAGCTTCCGGCGCACGAACTCGGGGCCGAGGAACACCGACGGCGGGGCGTGCTCCT 460

Qy   2184 AACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCG 2243
               |     ||| | ||   |||     |   || | || || | | | |||    |    ||
Db    459 CGTCGACCACGCTGTCCATCCAGTCGGACACGGGAACCTCCTCGATGACGCGGGTGCGCG 400

Qy   2244 CGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCT 2303
           || |  |||| ||  ||  | |      ||  |     |||| |   |||  | |    | |
Db    399 GGCCGTCGCAGTCGGGCATGAAGGCGGTGAGGAACACGGCGGCCGCGACCTTGTCGGGGA 340

Qy   2304 CGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCT 2363
             |      || |  |    |||    |    |||||| | || |        || |   | ||| |
Db    339 ACTCCTCGGCGGCGAGCGCGATGCTCATGCCGCCGAAGCTGTGGCCCACGAGCACCGCCC 280

Qy   2364 GAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAAC 2423
             |   |||||   ||| |    ||||  ||| | |  ||      |||  |||| ||  |||    ||
Db    279 GCTCGCCGTCCGGGAGGTCCCGGAGCGCGTCGAGCAGGGGGCGCGTGTAGTCCTCGAAGG 220

Qy   2424 CCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTAC 2483
              |  | ||  |     ||  |        || |   |          || |   | ||   ||
Db    219 TGGGCGCGTCGCTCAGCCGGCGCGCGTCGGAGCCGCATGCCGCGAGGTCCGGCGCGTCGA 160

Qy   2484 CCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGG 2543
            |         | ||   || |  |  |      ||| |     |      ||||| |   || |   |||
Db    159 CGCGGTGCCCCGCGGCGCGGAGGAGGGTGGCCACCTTGTACCAGCACCACCCGCCGTGGC 100

Qy   2544 C 2544
          |
Db     99 C 99


RESULT 3
CNS0091P/c
LOCUS       CNS0091P                 925 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL053013
VERSION     AL053013.1  GI:4934461
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
```

```
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..925
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR19D16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      120 a     61 c     61 g    172 t    511 others
ORIGIN

  Query Match              1.4%;  Score 46.4;  DB 17;  Length 925;
  Best Local Similarity   14.4%;  Pred. No. 0.1;
  Matches   42;  Conservative  133;  Mismatches  117;  Indels    0;  Gaps    0;

Qy       2185 ACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGC 2244
               :| ::::|||: :|   : : |:   | :   ||:     : ::::: ||:::::   :    : :
Db        837 MCSSSSSCCGSASARGVKVRASGGAGKRGGGSGGASASHSSSSACBSSSSSCSASCWSAS 778

Qy       2245 GCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTC 2304
              :: : ::::: :|     |:  ::::   ::::: :| :: : |:::   :   : :: : : :
Db        777 SSSSASSSRSRSGGGAGGSGASSSRSSSSSSSASAGSVVSSASSSSSSCSSSVSCSSVAS 718

Qy       2305 GCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTG 2364
              ::: :   :::    | :::: :       : :| | ::   :|:   |   :| :: :::::| :
Db        717 SMSCSSBSSSSASASSSSSSSSASCASCSCCCTSWSCSCSTSASMSAARSSSSSSSSCSS 658

Qy       2365 AAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACC 2424
                :: :: | :::|::::  |: ::    |||: :::|| : ::: :   : :::::      :
Db        657 SSMSASSSASSSSASSSSSSGSSSSSGSACGBSMSSGGGSGSVSASSGMSSSVSSSGGRS 598

Qy       2425 CAACAGGACGGTGGAACGTGGAACAGGGCGGTGTTGAAGGCAAAATGCGCG 2476
              :     || :|| |:   :| |::   : :: : : |    : : : |||    :|:
Db        597 SGSGGGGGVGGSGSSSSGSGSSGGSGSVCSCSSGCMCRCSCSSAAAAAASCV 546


RESULT 4
CNS006QP/c
LOCUS       CNS006QP               902 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR14J09 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL065804
VERSION     AL065804.1  GI:4944772
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 902)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..902
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR14J09"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      298 a     50 c     61 g    142 t    351 others
ORIGIN

  Query Match              1.4%;  Score 45.8;  DB 17;  Length 902;
  Best Local Similarity   16.5%;  Pred. No. 0.15;
  Matches   69;  Conservative  175;  Mismatches  173;  Indels    0;  Gaps    0;

Qy        120 CAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGC 179
              ::: :::: : :::: :  : ::::::::   ::::     ::::: |::::::|:    | :::
Db        888 VMVMVSRVAMMVMMVASACVCVSVVRSSSSSRSVSCSVSVVVGASVVVRVARVAGGRVSV 829

Qy        180 AGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGG 239
              |: : :   : :: :   :   ::::::::|      :: ::|:: ::::|   :    :     ::
Db        828 AVMRAVVVVASSRSAVRRAVVRVVMSRVARMVMMVSMVAMRMVVVVAVAVVRARMAAAVR 769

Qy        240 TCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCT 299
                 |:| :|         ::   : ::|:: ::|    ::::| ::|   |:|:||||::| :: :
Db        768 VAAGVAAVAAAAMAAVRMAMRVVAVRMVVAVARMMMMAVMVACMCMCMCCCVMCMMMMMM 709

Qy        300 GCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACC 359
              ::: :|:|:|::|     : ::: : :||   |   :: ::|| |:     ||  :: : |::| :
Db        708 VVMAMCSGRAVVCACAMAMRMAMAMAAMAAAAMVMMVAAAAMMMGGAVAMMAMVCMMARM 649

Qy        360 TGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGA 419
                |::    ||   :    :  : | |   :: :         : |             : |   |       | : |
Db        648 AACMMMMMAAMAMMAMAMMAAAMAMMAMAAMAMMMAACCCCCMCCMCCCCCCCCVMCMC 589

Qy        420 ATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGA 479
                |:   | : : :      ::      : : |:       :         | |::| :| ::    | :  ::: :
Db        588 CCCVACCMVAMMMCMCMMMCMAMAMMAMAMMAMAMMAAAAMMAAMAAMMAMCAMAMMMMM 529

Qy        480 AAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGG 536
              |::    || :::|: | |:||  :    :  :: |::::: : |::    :|::  :|||:| ||
Db        528 AMMMAACCMMMAMVAMAMGGGKAGDGKRKGKDRRRGRMCRVGVRAMDCVGGASGGGG 472


RESULT 5
AW707761
LOCUS       AW707761               471 bp    mRNA    linear   EST 01-MAY-2000
DEFINITION  832012E07.x1 C. reinhardtii CC-125 nutrient replete, -S, -Fe,
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   AW707761
VERSION     AW707761.1  GI:7593510
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
```

```
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
            Silflow,C., Stern,D. and Surzycki,R.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Elizabeth H. Harris
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000, USA
            Tel: 919 613 8164
            Fax: 919 613 8177
            Email: chlamy@duke.edu.
FEATURES             Location/Qualifiers
     source          1..471
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-125 wild type mt+ 137C"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-125 nutrient replete, -S,
                     -Fe, Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; This library combines cDNAs from nutrient replete
                     (TAP medium, mid-log phase), sulfur-deprived (TAP medium
                     -S for 2 hr), and iron-deprived cells from CC-125.
                     Fe-deprived cells were obtained by either (1) growth in
                     SGII medium at 0.18 micromolar Fe, 2X wash with SGII -Fe
                     and resuspension in SGII -Fe + 80 micromolar
                     bathophenanthroline sulfonic acid (an iron chelator) for 2
                     hours or (2) growth in SGII medium, 2X wash with SGII -Fe,
                     resuspension in SGII -Fe and continued growth for 48
                     hours. 1.25 micrograms of each polyA mRNA prep was used to
                     synthesize cDNA which was directionally cloned into lambda
                     ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage."
BASE COUNT       89 a    166 c    134 g     82 t
ORIGIN

  Query Match             1.4%;  Score 45.6;  DB 10;  Length 471;
  Best Local Similarity   48.5%;  Pred. No. 0.12;
  Matches  126;  Conservative    0;  Mismatches  134;  Indels    0;  Gaps    0;

Qy       2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
               | || |||| ||   |   ||||        || |   ||| |||   | |||| ||  |
Db        117 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 176

Qy       2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
              |      || ||  | |||  |  || |||   || |||   |   |  |||   |||| |    |
Db        177 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 236

Qy       3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
               ||||    ||       |       |     | |||      |||   ||  |  |  |  |  ||      | |
Db        237 CCGCATCCTGTTCGAGGTCGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 296

Qy       3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
               | ||||  |||||  |||      | |         | |        |||  | |      || ||||  |
Db        297 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 356

Qy       3174 CGGACGAGTCGGCGTAGGCT 3193
               |    |    || ||    ||||
Db        357 CACCCAGATCTGCTACGGCT 376


RESULT 6
CNS012S8
LOCUS       CNS012S8              1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN08I07 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL101954
VERSION     AL101954.1  GI:5613565
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN08I07"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT      253 a     65 c     73 g     53 t    657 others
ORIGIN

  Query Match             1.4%;  Score 44.8;  DB 17;  Length 1101;
  Best Local Similarity   10.5%;  Pred. No. 0.32;
  Matches   58;  Conservative  143;  Mismatches  346;  Indels    6;  Gaps    1;

Qy       2191 GCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCG 2250
              :::         :  |:                 |  :     :  :| :::|        ::       :
Db        306 SSSNNNNNNNASNAAVNNNNNNNNNNAAAANSNNNASNNSGNSSSASNNNNASSGNANNNS 365

Qy       2251 GCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCT 2310
              : |: :           :     :   |        :     |   :  :|        :  |      :  : ::::
Db        366 SAASASANNNNASNNNSANAANNNSSANAASASSSNSGGSNGNNSSAAGANSNSASSSSA 425

Qy       2311 ACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCC 2370
                :       |    :      |     :      ::       :::          |   :   :  :  |:  :  ||  ::
Db        426 SSNNAANASNSNNNAANNNSNANSSNNNSNSSVNNNSSNAANNSSNSNSNGSSSNAAASS 485

Qy       2371 GTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAG 2430
               :   :::|::|   :   ::|  |   ::   :              :       :     ::::|||  ::|  :  :
Db        486 SSSSSSSASSGNAASANSSAAANGSSSGSNNNNNNNSNNNNVNNASSSSAAANSVANSGS 545

Qy       2431 GACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACC 2490
              :  :|  ::||  :  :  :      |::  :      :  |          :          :    ::          : :
Db        546 SGASGSSVAAASASASSSNASSASANASSTNNNNNSNNNNSSNNSSTNNNNNSASNNNNG 605

Qy       2491 GTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCAT---- 2546
              :  :::|       ::          |    :  :  :                 |  |           :  |    ::  |
Db        606 SSNSSSANNNAVSNNNNASSNSNNSSNNNANNANNANNAAAANNNVAANGNSSNANNANS 665

Qy       2547 --GGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTG 2604
                  :|  :    |     |           |          |  :  :   ||    |    ||    ::|     :  |    |
Db        666 SSNSGGSNNAWNAANNNAANNNNSNSCSSSSGGAANTCNNAANNSSANNSAAAAANVNAA 725

Qy       2605 TTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTAC 2664
                  :  |    :        |   |      |::::::       ::::                    ::  :
Db        726 NNNSNAAASSGNNGATANNGAGSSSSSSNSSGSVSVNNVASNNNSNNSSNVSANANAANN 785

Qy       2665 GGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGC 2724
              :  :::        |    :  |              ::  ::  ::    :|:       :  :::               |    |     |::
Db        786 SNNSSSNNANANNVAANNNNANSSGSSSSSNNSGSNNGNSSSSVNNNNNAGNNANNAAVS 845

Qy       2725 GTCAACGGCACGC 2737
              :  |       ::::
```

```
Db    846 SNNANNNNAVSSS 858


RESULT 7
AL824948
LOCUS       AL824948                 584 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AL824948 p:234 Triticum aestivum cDNA clone A06_p234_plate_11, mRNA
            sequence.
ACCESSION   AL824948
VERSION     AL824948.1  GI:21836469
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
            Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
            Edwards,K.
  TITLE     A BBSRC-funded wheat EST resource for the academic community
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Barker G
            Institute of Arable Crop Research
            Long Ashton, Bristol BS41 9AF United Kingdom.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Triticum aestivum"
                     /cultivar="mercia"
                     /db_xref="taxon:4565"
                     /clone="A06_p234_plate_11"
                     /clone_lib="p:234"
                     /tissue_type="drought stressed seedlings"
                     /dev_stage="21 days old"
BASE COUNT      102 a    221 c    181 g     80 t
ORIGIN

  Query Match             1.4%;  Score 44.6;  DB 9;  Length 584;
  Best Local Similarity   56.5%;  Pred. No. 0.27;
  Matches   83;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy       1219 ATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAG 1278
Db        351 ATCTCCGGGACGTCCATGTCGTGCCCGCACGTGAGCGGCCTCGCGGCGCTGCTCCGGCAG 410

Qy       1279 AAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGAC 1338
Db        411 GCGCACCCGGAGTGGAGCCCCGCGGCGGTCAAGTCGGCGCTCATGACCACGGCGTACAAC 470

Qy       1339 ATCGGTGCAGTCGGCGTGGACAGCAAG 1365
Db        471 CTGGACAACTCCGGCGAGATCATCAAG 497


RESULT 8
BE726719
LOCUS       BE726719                 488 bp    mRNA    linear   EST 14-SEP-2000
DEFINITION  894093B05.y3 C. reinhardtii CC-1690, normalized, Lambda Zap II
            Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BE726719
VERSION     BE726719.1  GI:10127903
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
            McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants; project phase 2
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
                     II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; This library, constructed by John Davies and Jeffrey
                     McDermott, combines cDNAs from CC-1690 cells grown to
                     mid-log phase in TAP (acetate-containing) medium in the
                     light, TAP medium in the dark, HS (minimal) medium in
                     ambient levels of CO2 and HS medium bubbled with 5% CO2.
                     PolyA mRNA was purified from each sample, pooled and cDNA
                     synthesized. The cDNA was directionally cloned into lambda
                     ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage. The library was normalized using method 4 described
                     in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT       88 a    177 c    140 g     83 t
ORIGIN

  Query Match             1.4%;  Score 44;  DB 12;  Length 488;
  Best Local Similarity   48.1%;  Pred. No. 0.37;
  Matches  125;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy       2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
Db        103 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 162

Qy       2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
Db        163 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 222

Qy       3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
Db        223 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 282

Qy       3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
Db        283 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 342

Qy       3174 CGGACGAGTCGGCGTAGGCT 3193
Db        343 CACCCAGATCTGCTACGGCT 362


RESULT 9
AV639324
LOCUS       AV639324                 490 bp    mRNA    linear   EST 15-DEC-2000
DEFINITION  AV639324 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC097h07_r 5', mRNA sequence.
ACCESSION   AV639324
VERSION     AV639324.1  GI:10782652
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 490)
  AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
  TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
  JOURNAL   DNA Res. 7 (5), 305-307 (2000)
```

```
  MEDLINE   20539644
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1..490
                     /organism="Chlamydomonas reinhardtii"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="HC097h07_r"
                     /clone_lib="Chlamydomonas reinhardtii 5% CO2"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from cells cultured
                     in a medium with bubbling air containing 5% carbon
                     dioxide"
BASE COUNT       89 a    176 c    142 g     83 t
ORIGIN

  Query Match             1.4%;  Score 44;  DB 10;  Length 490;
  Best Local Similarity   48.1%;  Pred. No. 0.37;
  Matches  125;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy       2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
               | || |||| ||    |  ||||        || |  ||| |||  | |||| ||  |
Db         98 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 157

Qy       2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
              |      || ||  | ||| |  || |||  || |||  |   | |||  |||| |    |
Db        158 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 217

Qy       3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
                ||||    ||      |          |    | |||    |||  || | | | ||       | |
Db        218 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 277

Qy       3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
              | ||||  |||||  |||        | |        |||  | |      || ||||   |
Db        278 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 337

Qy       3174 CGGACGAGTCGGCGTAGGCT 3193
              |    |    || ||    ||||
Db        338 CACCCAGATCTGCTACGGCT 357


RESULT 10
AV636203
LOCUS       AV636203                 497 bp    mRNA    linear   EST 15-DEC-2000
DEFINITION  AV636203 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC056e01_r 5', mRNA sequence.
ACCESSION   AV636203
VERSION     AV636203.1  GI:10779523
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
  TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
  JOURNAL   DNA Res. 7 (5), 305-307 (2000)
  MEDLINE   20539644
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1..497
                     /organism="Chlamydomonas reinhardtii"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="HC056e01_r"
                     /clone_lib="Chlamydomonas reinhardtii 5% CO2"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from cells cultured
                     in a medium with bubbling air containing 5% carbon
                     dioxide"
BASE COUNT       90 a    179 c    144 g     84 t
ORIGIN

  Query Match             1.4%;  Score 44;  DB 10;  Length 497;
  Best Local Similarity   48.1%;  Pred. No. 0.37;
  Matches  125;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy       2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
               | || |||| ||    |  ||||        || |  ||| |||  | |||| ||  |
Db        108 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 167

Qy       2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
              |      || ||  | ||| |  || |||  || |||  |   | |||  |||| |    |
Db        168 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 227

Qy       3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
                ||||    ||      |          |    | |||    |||  || | | | ||       | |
Db        228 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 287

Qy       3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
              | ||||  |||||  |||        | |        |||  | |      || ||||   |
Db        288 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 347

Qy       3174 CGGACGAGTCGGCGTAGGCT 3193
              |    |    || ||    ||||
Db        348 CACCCAGATCTGCTACGGCT 367


RESULT 11
AV636123
LOCUS       AV636123                 513 bp    mRNA    linear   EST 15-DEC-2000
DEFINITION  AV636123 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC055d12_r 5', mRNA sequence.
ACCESSION   AV636123
VERSION     AV636123.1  GI:10779443
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
  TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
  JOURNAL   DNA Res. 7 (5), 305-307 (2000)
  MEDLINE   20539644
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1..513
                     /organism="Chlamydomonas reinhardtii"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="HC055d12_r"
                     /clone_lib="Chlamydomonas reinhardtii 5% CO2"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from cells cultured
                     in a medium with bubbling air containing 5% carbon
                     dioxide"
BASE COUNT       97 a    181 c    149 g     86 t
ORIGIN
```

```
  Query Match             1.4%;  Score 44;  DB 10;  Length 513;
  Best Local Similarity   48.1%;  Pred. No. 0.38;
  Matches  125;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy     2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
            | || |||| ||   |  ||||        || |  ||| |||  | |||| ||  |
Db      245 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 304

Qy     2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
            |       || ||  | ||| |  || |||  || |||  |  | |||  |||| |    |
Db      305 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 364

Qy     3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
             ||||   ||    |         |   |  |||   |||  || | | | ||        | |
Db      365 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 424

Qy     3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
            | ||||  |||||  |||        | |       | |     |||  | |     || ||||  |
Db      425 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 484

Qy     3174 CGGACGAGTCGGCGTAGGCT 3193
            |  |   || ||   ||||
Db      485 CACCCAGATCTGCTACGGCT 504
```

```
RESULT 12
BQ822928
LOCUS       BQ822928                 557 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  1030104H10.y1 C. reinhardtii CC-1690, Deflagellation (normalized),
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BQ822928
VERSION     BQ822928.1  GI:22073590
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
            ,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1030
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Deflagellation
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Deflagellation library, constructed by John Davies
                     and Jeffrey McDermott, combines cDNAs from CC-1690 cells
                     which had been re-synthesizing flagella for 15, 30 and 60
                     min after being deflagellated by pH shock. PolyA mRNA was
                     purified from each sample, pooled and cDNA synthesized.
                     The cDNA was directionally cloned into lambda Zap II
                     (Stratagene) in the EcoRI (5') and XhoRI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage. The library was normalized using method 4 described
                     in Bonaldo et al., (1996) Genome Research 6: 791-806."
BASE COUNT       92 a    188 c    168 g    109 t
ORIGIN
```

```
  Query Match             1.4%;  Score 44;  DB 14;  Length 557;
  Best Local Similarity   48.1%;  Pred. No. 0.39;
  Matches  125;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy     2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
            | || |||| ||   |  ||||        || |  ||| |||  | |||| ||  |
Db       95 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 154

Qy     2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
            |       || ||  | ||| |  || |||  || |||  |  | |||  |||| |    |
Db      155 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 214

Qy     3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
             ||||   ||    |         |   |  |||   |||  || | | | ||        | |
Db      215 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 274

Qy     3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
            | ||||  |||||  |||        | |       | |     |||  | |     || ||||  |
Db      275 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 334

Qy     3174 CGGACGAGTCGGCGTAGGCT 3193
            |  |   || ||   ||||
Db      335 CACCCAGATCTGCTACGGCT 354
```

```
RESULT 13
AV387616
LOCUS       AV387616                 562 bp    mRNA    linear   EST 29-SEP-2000
DEFINITION  AV387616 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
            cDNA clone CM027a08_r, mRNA sequence.
ACCESSION   AV387616
VERSION     AV387616.1  GI:6541832
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S., Fukuzawa,H. and Tabata,S.
  TITLE     A large scale structural analysis of cDNAs in a unicellular green
            alga, Chlamydomonas reinhardtii. I. Generation of 3433
            non-redundant expressed sequence tags
  JOURNAL   DNA Res. 6 (6), 369-373 (1999)
  MEDLINE   20152988
COMMENT     Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .562
                     /organism="Chlamydomonas reinhardtii"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="CM027a08_r"
                     /clone_lib="Chlamydomonas reinhardtii C9"
                     /dev_stage="photoautotrophic growth"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT       97 a    186 c    173 g    106 t
ORIGIN
```

```
  Query Match             1.4%;  Score 44;  DB 10;  Length 562;
  Best Local Similarity   48.1%;  Pred. No. 0.39;
  Matches  125;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy     2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
            | || |||| ||   |  ||||        || |  ||| |||  | |||| ||  |
Db       95 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 154

Qy     2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
            |       || ||  | ||| |  || |||  || |||  |  | |||  |||| |    |
Db      155 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 214
```

```
Qy     3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
             ||||  ||      |          |   | |||   |||  || | | | ||     ||
Db      215 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 274

Qy     3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
             | ||||  |||||  |||      | |        | |    |||  | |     || ||||  |
Db      275 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 334

Qy     3174 CGGACGAGTCGGCGTAGGCT 3193
             |   |   || ||    ||||
Db      335 CACCCAGATCTGCTACGGCT 354


RESULT 14
BI875064
LOCUS       BI875064                 634 bp    mRNA    linear   EST 11-OCT-2001
DEFINITION  963121E11.y1 C. reinhardtii CC-1690, Stress condition I, normalized
            , Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BI875064
VERSION     BI875064.1  GI:16073068
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
            ,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 963
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1..634
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Stress condition I,
                     normalized, Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; This library, constructed by John Davies and Jeffrey
                     McDermott, combines cDNAs from CC-1690 cells grown to
                     mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
                     1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr
                     , 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
                     purified from each sample, pooled and cDNA synthesized.
                     The cDNA was directionally cloned into lambda Zap II
                     (Stratagene) in the EcoRI (5') and XhoRI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage. The library was normalized using method 4 described
                     in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT      114 a     226 c     185 g     109 t
ORIGIN

  Query Match             1.4%;  Score 44;  DB 13;  Length 634;
  Best Local Similarity   48.1%;  Pred. No. 0.41;
  Matches  125;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy     2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
             | || |||| ||    |   ||||        || |   ||| |||   | |||| ||  |
Db      315 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 374

Qy     2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
             |      || ||  | ||| |  || |||  || |||   |   | |||   |||| |    |
```

```
Db      375 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 434

Qy     3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
             ||||  ||      |          |   | |||   |||  || | | | ||     | |
Db      435 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 494

Qy     3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
             | ||||  |||||  |||      | |        | |    |||  | |     || ||||  |
Db      495 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGACCAGCG 554

Qy     3174 CGGACGAGTCGGCGTAGGCT 3193
             |   |   || ||    ||||
Db      555 CACCCAGATCTGCTACGGCT 574


RESULT 15
CNS0106Y
LOCUS       CNS0106Y                1087 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN03J01 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL098596
VERSION     AL098596.1  GI:5610207
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1087)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1087
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN03J01"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT      236 a     236 c     241 g     130 t     244 others
ORIGIN

  Query Match             1.4%;  Score 43.8;  DB 17;  Length 1087;
  Best Local Similarity   28.6%;  Pred. No. 0.61;
  Matches   54;  Conservative   62;  Mismatches   73;  Indels    0;  Gaps    0;

Qy     2385 GGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTG 2444
             ||:| :| |||: |||          :  | |   |    |||::::  ||:| |:|  :    :: :
Db      720 GGMCARCCACGMAACGAAGGGMSAAGCCTRCCAACAAMMMMGCCARGSCSGVVACCSSVS 779

Qy     2445 GGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGC 2504
             ||:::    :::| |   |||:|:|: : ||:|||  :  : :::   |||   |:|:  :|| :
Db      780 GGMMMCCCSSSGAGRAGAASGBAVGVCGCSCGGRMMAVAMSMCCCCGRAGSCRSGSCCAM 839

Qy     2505 GAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAG 2564
             |:  ::|:::::|      ::   ::   : ::|:   :     : :          |: ::: | ||::
Db      840 GRGSRSCSVMSVASCGCSRGGMSGCSKSVCSGGMAAMSCSCSACCSGASGRMRGACGGVS 899

Qy     2565 CGAAAACAG 2573
             | |:: ||:
Db      900 CMARRSCAS 908
```

```
RESULT 16
CNS017X5
LOCUS       CNS017X5              1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN37L12 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL108611
VERSION     AL108611.1  GI:5628915
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN37L12"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      273 a    245 c    164 g     94 t    325 others
ORIGIN

  Query Match             1.4%;  Score 43.8;  DB 17;  Length 1101;
  Best Local Similarity   22.5%;  Pred. No. 0.61;
  Matches   71;  Conservative  115;  Mismatches  128;  Indels    1;  Gaps    1;

Qy       2048 CACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAAC 2107
              ||    |: ::   | :|:   :::|:|   : |:|:::   :     | | |:|||:: |   : ::
Db        779 CAACAAMRSMAGCAVCMMARVMAMACVMAGVGVMMCAVCAMCGGCAMCACSSGGGGRGVV 838

Qy       2108 AGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAA 2167
              ::|| ||       || | :|::|       | | ::   :|   :|      | : |: |   ||:
Db        839 MRGGGGGGMAAASCASARAMRCAAASGAGRCCMMAARGGCRCACARGGVRCMRCAMCAM 898

Qy       2168 CACCCGAGAC-GGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCC 2226
              :::      :||   :   :|  ::  ::::: :|: ::|  ::      :  :   | ::|:: |:   | |
Db        899 MMMAGCSRACAASVMSACMMRRSRSMGSCMMMMCMMMACARVRCACARRGSVRCMCACRC 958

Qy       2227 TACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTA 2286
               :|::::: ::     :|: :  ||| ::||::: ::: |: ::::   |: |     :| : : |
Db        959 MVCRSVVSSVSACCMCRASGAGCASSAGMSRCMMMMCMRRSRRGGGSGGGACVCASMVMA 1018

Qy       2287 CGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGAT 2346
              ::|| :   :||::       :|  :  :  |::                |:|::   :         ::|::   :   |
Db       1019 VVCARMAMMAAVMASGRSGRSCMCAMMACAMCAMACVAMRCVVAGCAASMCMMVMMACAC 1078

Qy       2347 ATGCAGGGACGCCGG 2361
              : |||::  ||  ::::
Db       1079 MMGCASRCACASVRS 1093


RESULT 17
AW676029
LOCUS       AW676029               655 bp    mRNA    linear   EST 01-MAY-2000
DEFINITION  832001E09.y1 C. reinhardtii CC-125 nutrient replete, -S, -Fe,
            Lambda Zap II Chlamydomonas reinhardtii cDNA similar to Sbp2 gene
            encoding seduheptulose-1,7-bisphosphatase, mRNA sequence.
ACCESSION   AW676029
VERSION     AW676029.1  GI:7545541
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
            Silflow,C., Stern,D. and Surzycki,R.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Elizabeth H. Harris
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000, USA
            Tel: 919 613 8164
            Fax: 919 613 8177
            Email: chlamy@duke.edu
            Sbp2 (Csbp) gene; compare C. reinhardtii Accession X74418.
FEATURES             Location/Qualifiers
     source          1..655
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-125 wild type mt+ 137C"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-125 nutrient replete, -S,
                     -Fe, Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; This library combines cDNAs from nutrient replete
                     (TAP medium, mid-log phase), sulfur-deprived (TAP medium
                     -S for 2 hr), and iron-deprived cells from CC-125.
                     Fe-deprived cells were obtained by either (1) growth in
                     SGII medium at 0.18 micromolar Fe, 2X wash with SGII -Fe
                     and resuspension in SGII -Fe + 80 micromolar
                     bathophenanthroline sulfonic acid (an iron chelator) for 2
                     hours or (2) growth in SGII medium, 2X wash with SGII -Fe,
                     resuspension in SGII -Fe and continued growth for 48
                     hours. 1.25 micrograms of each polyA mRNA prep was used to
                     synthesize cDNA which was directionally cloned into lambda
                     ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage."
BASE COUNT      117 a    234 c    192 g    112 t
ORIGIN

  Query Match             1.3%;  Score 43.2;  DB 10;  Length 655;
  Best Local Similarity   49.6%;  Pred. No. 0.72;
  Matches  111;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy       2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
              | || |||| ||    |   ||||          || |   ||| |||   | |||| ||   |
Db        289 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 348

Qy       2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
              |        || ||   | ||| |  || |||   || |||   |   | |||   ||||  |     |
Db        349 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 408

Qy       3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
                 ||||    ||      |              |     |  |||    |||   || |  |  |  | ||          | |
Db        409 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 468

Qy       3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAAC 3157
              | ||||   |||||   ||    |||    | |||    |          | || |  |
Db        469 CAGCTGCGACGGCAAGGGCCGTGTCCGCTCTGGACATCCCCATC 512


RESULT 18
```

```
AV633578
LOCUS       AV633578                477 bp    mRNA    linear   EST 15-DEC-2000
DEFINITION  AV633578 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC022e07_r 5', mRNA sequence.
ACCESSION   AV633578
VERSION     AV633578.1  GI:10776898
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
  TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
  JOURNAL   DNA Res. 7 (5), 305-307 (2000)
  MEDLINE   20539644
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Chlamydomonas reinhardtii"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="HC022e07_r"
                     /clone_lib="Chlamydomonas reinhardtii 5% CO2"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from cells cultured
                     in a medium with bubbling air containing 5% carbon
                     dioxide"
BASE COUNT       87 a    168 c    138 g     83 t      1 others
ORIGIN

  Query Match             1.3%;  Score 43;  DB 10;  Length 477;
  Best Local Similarity   47.7%;  Pred. No. 0.71;
  Matches  124;  Conservative    0;  Mismatches  136;  Indels    0;  Gaps    0;

Qy     2934 CTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACG 2993
             | || |||| ||    |  ||||       || |  ||| |||  | |||| ||  |
Db      211 CCTGGGCGAGAAGTACACCCTGCGCTACACCGGCGGCATGGTGCCCGACGTGTTCCAGAT 270

Qy     2994 CGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGC 3053
             |      || ||  | ||| |  || |||  || |||  |  | |||  |||| |   |
Db      271 CATTGTCAAGGAGAAGGGTGTGTTCACCAACGTGACCTCGCCCACCACCAAGGCCAAGCT 330

Qy     3054 ACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAA 3113
              ||||  ||     |       |  | |||   |||  || | | | | ||     | |
Db      331 CCGCATCCTGTTCGAGGTTGCTCCCCTGGCCCTGCTGATCGAGAAGGCCGGCGGTGCCTC 390

Qy     3114 CGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAG 3173
             | ||||  |||||  |||     | |      | |     |||  | |    || | | || |
Db      391 CAGCTGCGACGGCAAGGCCGTGTCCGCTCTGGACATCCCCATCCTGGTGTGCGANCAGCG 450

Qy     3174 CGGACGAGTCGGCGTAGGCT 3193
             |   |    || ||    ||||
Db      451 CACCCAGATCTGCTACGGCT 470


RESULT 19
BG947100
LOCUS       BG947100                358 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  IP1_2_E12.b1_A002 Immature pannicle 1 (IP1) Sorghum bicolor cDNA,
            mRNA sequence.
ACCESSION   BG947100
VERSION     BG947100.1  GI:14365289
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 358)
  AUTHORS   Klein,R.R., Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: developing preanthesis pannicles
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for high quality sequence is
            20. Three-prime sequences, which are obtained with PolyTMix or T7
            sequencing primer, are presented as the reverse complement.
            Seq primer: JEN REV
            High quality sequence stop: 352
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .358
                     /organism="Sorghum bicolor"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone_lib="Immature pannicle 1 (IP1)"
                     /note="Organ: Developing preanthesis pannicles; Vector:
                     pBluescript II SK(-) from Lambda Zap II; Site_1: XhoI;
                     Site_2: EcoRI; The library was made from poly-A RNA in the
                     cloning vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT       55 a    129 c    128 g     46 t
ORIGIN

  Query Match             1.3%;  Score 42.2;  DB 13;  Length 358;
  Best Local Similarity   55.0%;  Pred. No. 1;
  Matches   83;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy     1219 ATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAG 1278
             ||  |||| || ||| | ||      ||    ||| | |||     ||||| ||||| |    |
Db       24 ATCTCCGGGACGTCCATGTCGTGCCCGCACGTGAGCGGGCTTGCGGCGCTGCTCCGCGCG 83

Qy     1279 AAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGAC 1338
                 |||||  |  ||||   |     | | ||  |  ||||| |||| |||||  |  ||
Db       84 GCGCACCCGGAGTGGAGCCCGGCAGCCATCCGGTCGGCGCTGATGACGACGGCGTACAAC 143

Qy     1339 ATCGGTGCAGTCGGCGTGGACAGCAAGTTCG 1369
                 ||| |||||  | || |   |   ||
Db      144 GATTACCCAGGCGGCGCTGGCATCCTGGACG 174


RESULT 20
BE041076/c
LOCUS       BE041076                755 bp    mRNA    linear   EST 07-JUN-2000
DEFINITION  OF19D09 OF Oryza sativa cDNA 5' similar to glycine-rich protein,
            mRNA sequence.
ACCESSION   BE041076
VERSION     BE041076.1  GI:8336790
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 755)
  AUTHORS   Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrea
            ,H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C.,
            Scara,G., Wheeler,M. and Zepeda,G.R.
  TITLE     Functional Genomics of Plant Stress Tolerance
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Michalowski,C.B.
            University of Arizona
```

```
            Bio Sciences West room 513, Tucson, AZ 85721, USA
            Tel: 520-621-7982
            Fax: 520-621-1697
            Email: cbm@u.arizona.edu
            An open reading frame exists.
FEATURES             Location/Qualifiers
     source          1..755
                     /organism="Oryza sativa"
                     /strain="Pokkali"
                     /db_xref="taxon:4530"
                     /clone_lib="OF"
                     /tissue_type="entire plant"
                     /dev_stage="2 weeks"
                     /note="1 week 150mM NaCl"
BASE COUNT      105 a    216 c    258 g    176 t
ORIGIN

  Query Match              1.3%;  Score 42.2;  DB 10;  Length 755;
  Best Local Similarity   47.5%;  Pred. No. 1.5;
  Matches  125;  Conservative    0;  Mismatches  138;  Indels    0;  Gaps    0;

Qy       2109 GGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAAC 2168
              || |  || ||||| ||     || | |    ||        |  ||| | ||| |  || | ||
Db        515 GGCCCCCACCAATCAGGTCCTCCAGGTCCCCGCCGGAGTCGCCACGGAAGCCACCACCAC 456

Qy       2169 ACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTA 2228
                | |  || || |    | || ||||||| |     | |            | |  |  | ||
Db        455 GACTGCCACCGTAGCCGCCGCCGCCGCCGCCACGGCCGCAGCAGTAGCCGCCGCCGCCAC 396

Qy       2229 CGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACG 2288
              ||  ||  |  ||    ||  | | |||||               |     |||  | ||   |
Db        395 CGTAGCCGCCCTCACGGCGCCCGTAGCCGCCGCCGCCACCTCCTCCGTAGCCGCCGCCTC 336

Qy       2289 CATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATAT 2348
              ||     |  |     | ||||| |||| |    ||||         ||||| ||    ||||
Db        335 CACGACCGCCGCCGTAGCCGCCACCGCCGCCGCCGTAGCCCCCGCCTCCGCCGCCGGAGC 276

Qy       2349 GCAGGGACGCCGGCTGAAAGCCG 2371
              |  |||||   | || |  | ||
Db        275 GGCGGGACTGGGCCTCATTGACG 253


RESULT 21
CNS017GS/c
LOCUS       CNS017GS             1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN17J04 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL108022
VERSION     AL108022.1  GI:5628326
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN17J04"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      268 a    126 c    100 g    195 t    412 others
ORIGIN

  Query Match              1.3%;  Score 42;  DB 17;  Length 1101;
  Best Local Similarity   19.3%;  Pred. No. 2;
  Matches   80;  Conservative  148;  Mismatches  187;  Indels    0;  Gaps    0;

Qy        254 CGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACT 313
              |::  : ||| : | : ::: :::  :|  :  :: :      |:   |     | :|:|::
Db       1087 CVSSVMGAGAGRCGMRRVVMSRRMAAVAVAMAGMMAMAAAAAMVRRGAGVGCRVGMGVSS 1028

Qy        314 TTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAG 373
                ::  :|:::::  :| |: : :: :        | ::::| :|||::  :|:   :|:::
Db       1027 VGMSGSRCSSMRVVCRCMCRCMAMMMVSMRAAAAVRMMCAMAAAMMMAMAMAAAVARMRM 968

Qy        374 GCTATACAGGACGCGGGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCA 433
                  | |:    : | |::: | ||: ::     |: :: :: ::: ::: ::: : ::: :|
Db        967 AAAAMASGAARGGRGVRRGAMAGSSVSAGCGCRVVRCBVRMVSASMMMMVMMCSMMAMA 908

Qy        434 TATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACT 493
                : ::   :::::|:  : :  :: :|: |:::  :|  : : ::::     : |||||
Db        907 AMAMMMAAMMVVMASAAWTYWYHMARARGGRRVGRRGRRDVRSSVVRGGRGRGAAAAAGA 848

Qy        494 ATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTT 553
              | ::: |  :         :||  :| |::|   : :|  |:       |:| |      |::
Db        847 ACMVVCCCVVVGRAAARAAAARAACVSCVCCMCMCCAAVMAMMCAMAAARSGRSGARMSS 788

Qy        554 CTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAA 613
                :|      :: : :: ||  |  :: |  :      | |  : :::|:| :|   :   : :||:
Db        787 ASHTWTTDBRASGSSACGAAAACRVMARSVGGACCASCMGRRSGRAGWTTTKDDDWDTAR 728

Qy        614 AAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGA 668
              :| |||| :    :   :   :|       |   | ::|   : ||   :  :|:||: :| :|
Db        727 RAMAAATRKWNWRATWTTWTTGTTTAACAMWTTRKTTTTTMAASCVTTVTRTADA 673


RESULT 22
CNS017SY
LOCUS       CNS017SY             1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL108460
VERSION     AL108460.1  GI:5628764
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
```

```
                    /clone="BACN37L08"
                    /clone_lib="DrosBAC"
                    /plasmid="pBeloBAC11"
                    /note="end : SP6"
BASE COUNT      254 a    176 c    160 g    152 t    359 others
ORIGIN

  Query Match            1.3%;  Score 42;  DB 17;  Length 1101;
  Best Local Similarity  16.2%;  Pred. No. 2;
  Matches   56;  Conservative  157;  Mismatches  131;  Indels    2;  Gaps    1;

Qy       2242 CGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGT 2301
              ::::|| :|:   |   :    |: | | : : | :|||  :|: | ||::| |   : |::
Db        738 MVSSCACSGSGSGGASACGASGAGGGSGRGGGSGCCASSGVCGAACSSASC--SSASMCG 795

Qy       2302 CTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGG 2361
              : :|:: :   | : | : :|: ::|: |::::::: :|   :|:| | |:   | ::::|::
Db        796 VSSGSSCSASCGSCCGVSSCSAVSASSASSVMSKVASAVASCSAVASGMSAGAVSSSCRS 855

Qy       2362 CTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAA 2421
              :     ::    : ::: ::::  ::    |: |::: |:: :            : :::: :
Db        856 SVASSVSAAASVSSSVSSSSSSSVVSAAAVSASSSSASSASMAVAAAAAVAVSSVSVASV 915

Qy       2422 ACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGT 2481
              ::::  | :: : : :|: |::  :    : :::::: :| : |::: : | ::: ::: :
Db        916 VSSSSSCSSSSSASVVVSASVAASASSVSSSSSSSVSTSSSASVSSVSAVSMSAVVSSSA 975

Qy       2482 ACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTG 2541
               ::: : ::: : : |            |||| ::::  |:| ::   : ::   ::::: :|: :
Db        976 SSSSSVSVSSVVAVAASAAAAAAAAAAAASSSSASAVAVVSASSSSSSASSSSSSSSASSS 1035

Qy       2542 GGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAAAA 2587
              :::: ::: ::: :: :   ::: :   | |:|: :::|| : ||:||
Db       1036 VSSVSSSSCSVVSVVSSSSVSVAVASASASASVSSVAAAVAAAMAA 1081


RESULT 23
AL716352
LOCUS       AL716352                 536 bp    mRNA    linear   EST 18-APR-2002
DEFINITION  AL716352 Danio rerio embryonic inner ear subtracted cDNA Danio
            rerio cDNA clone BN0AA016ZC03 5', mRNA sequence.
ACCESSION   AL716352
VERSION     AL716352.1  GI:20180955
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin
            ,J.P., Weissenbach,J. and Petit,C.
  TITLE     A subtracted cDNA library from the zebrafish (Danio rerio)
            embryonic inner ear
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="BN0AA016ZC03"
                     /clone_lib="Danio rerio embryonic inner ear subtracted
                     cDNA"
                     /tissue_type="inner ear"
                     /dev_stage="embryonic"
                     /note="subtracted cDNA library"
BASE COUNT      175 a    106 c    125 g    130 t
ORIGIN

  Query Match            1.3%;  Score 41.8;  DB 9;  Length 536;
  Best Local Similarity  51.3%;  Pred. No. 1.7;
  Matches   97;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy        463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
               | | | ||  | ||  ||| | | |      | ||    |  | |||| |||    | |  ||
Db        331 GTAAATGGTGAAAAGAAAAGTGAAAGTTCATCTAGTTCAGAAATGCAGAATACATCATCT 390

Qy        523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
                || | |   |   |   || |||   ||||||   || | ||   |    | ||  |||| |
Db        391 AAACAGGTTGTAGTGGAACACAAAGAAGCTTTGATCCAAGAAAACAGCATTGAAGAGGAT 450

Qy        583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
              |||  || |    |  ||      | |    | |||| |||||   | |||  |        ||  |
Db        451 GAACAAATGGATAAAGAAGAGAGTCCATTTAAAGGAATCGCCCCCATAAAGACTTTCAGC 510

Qy        643 CATATTATT 651
              | |||| ||
Db        511 CCTATTGTT 519


RESULT 24
BQ283304/c
LOCUS       BQ283304                 538 bp    mRNA    linear   EST 13-MAY-2002
DEFINITION  WHE3088_D10_G20ZS Wheat cold-stressed seedling subtracted cDNA
            library Triticum aestivum cDNA clone WHE3088_D10_G20, mRNA
            sequence.
ACCESSION   BQ283304
VERSION     BQ283304.1  GI:20552834
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Crossman,C., Fenton
            ,R.D., Lazo,G.R., Nguyen,H.T., Pham,J., Rausch,C.J., Wilson,C., Woo
            ,J. and Zhang,D.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Cold-stressed seedling subtracted cDNA library
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE3088_D10_G20"
                     /clone_lib="Wheat cold-stressed seedling subtracted cDNA
                     library"
                     /tissue_type="Seedling"
                     /dev_stage="Five-day old seedling"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
                     ,germinated and grown aseptically in the dark at room
                     temperature on filter paper with water,nystatin and
                     cefotaxime in covered crystallization dishes. Five-day old
                     seedlings were transferred to 5 C cold room and kept for
                     48 hr. The tissue, total RNA, and poly(A) RNA were
                     prepared, a cDNA library was made,and the cDNA clones were
                     in vivo excised to give pBluescript phagemids in the TJ
                     Close lab (Choi,Close,Fenton) at the University of
                     California,Riverside. The cDNA clones were in vivo excised
```

```
                 to give pBluescript phagemids before subtraction was
                 carried out. The mass excision of phagemid library and
                 subtraction were done in HT Nguyen lab by D. Zhang at
                 Texas Tech University. Plasmid DNA preparations and DNA
                 sequencing were performed in the OD Anderson lab (all
                 other authors)."
BASE COUNT      109 a    198 c    145 g     86 t
ORIGIN

  Query Match             1.3%;  Score 41.8;  DB 14;  Length 538;
  Best Local Similarity   54.1%;  Pred. No. 1.7;
  Matches   85;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          | | ||| || |  | ||| | |||||| ||||||  | |||  |    |||    |||  ||
Db    397 GTTGAAGGCGGCGCGGTTGGATCGGCGGACGTGGAGGGAGACGTACTCGGCACCGAAGAC 338

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
            |||    |||||| |   |||||     ||          ||| |||| | |
Db    337 CTGGTCCATGGCGGCCTGCGCGGCGGACATGAGCTTGGTGGCGAGACCGAGCTTGCGGTG 278

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGG 3097
          |    || |    ||||| ||   |   | | ||| || ||
Db    277 AGAGCGGAGGACCGCCAGGGAGGTGATGTGGCCGTGG 241


RESULT 25
AL721811
LOCUS       AL721811                 650 bp    mRNA    linear   EST 18-APR-2002
DEFINITION  AL721811 Danio rerio embryonic inner ear subtracted cDNA Danio
            rerio cDNA clone BN0AA054ZA05 3', mRNA sequence.
ACCESSION   AL721811
VERSION     AL721811.1  GI:20186415
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 650)
  AUTHORS   Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin
            ,J.P., Weissenbach,J. and Petit,C.
  TITLE     A subtracted cDNA library from the zebrafish (Danio rerio)
            embryonic inner ear
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="BN0AA054ZA05"
                     /clone_lib="Danio rerio embryonic inner ear subtracted
                     cDNA"
                     /tissue_type="inner ear"
                     /dev_stage="embryonic"
                     /note="subtracted cDNA library"
BASE COUNT      213 a    122 c    160 g    155 t
ORIGIN

  Query Match             1.3%;  Score 41.8;  DB 9;  Length 650;
  Best Local Similarity   51.3%;  Pred. No. 1.8;
  Matches   97;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy    463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
          | | | | ||  | ||  ||| | | |       | ||    |  | |||| ||| | | |  ||
Db    331 GTAAATGGTGAAAAGAAAAGTGAAAGTTCATCTAGTTCAGAAATGCAGAATGCATCATCT 390

Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            || | |   |   |   || |||   ||||||   || | ||   |     | ||  |||| |
Db    391 AAACAGGTTGTAGTGGAACACAAAGAAGCTTTGATCCAAGAAAACAGCATTGAAGAGGAT 450

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
          |||  || |     |  ||        | |    | |||| ||||| | |||  |        |  |
Db    451 GAACAAATGGATAAAGAAGAGAGTCCATTTAAAGGAATCGCTCCCATAAAGACTTCCAGC 510

Qy    643 CATATTATT 651
          | |||| ||
Db    511 CCTATTGTT 519


RESULT 26
AL721550
LOCUS       AL721550                 766 bp    mRNA    linear   EST 18-APR-2002
DEFINITION  AL721550 Danio rerio embryonic inner ear subtracted cDNA Danio
            rerio cDNA clone BN0AA053ZA12 3', mRNA sequence.
ACCESSION   AL721550
VERSION     AL721550.1  GI:20186154
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 766)
  AUTHORS   Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin
            ,J.P., Weissenbach,J. and Petit,C.
  TITLE     A subtracted cDNA library from the zebrafish (Danio rerio)
            embryonic inner ear
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .766
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="BN0AA053ZA12"
                     /clone_lib="Danio rerio embryonic inner ear subtracted
                     cDNA"
                     /tissue_type="inner ear"
                     /dev_stage="embryonic"
                     /note="subtracted cDNA library"
BASE COUNT      243 a    147 c    194 g    182 t
ORIGIN

  Query Match             1.3%;  Score 41.8;  DB 9;  Length 766;
  Best Local Similarity   51.3%;  Pred. No. 2;
  Matches   97;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy    463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
          | | | | ||  | ||  ||| | | |       | ||    |  | |||| ||| | | |  ||
Db    376 GTAAATGGTGAAAAGAAAAGTGAAAGTTCATCTAGTTCAGAAATGCAGAATGCATCATCT 435

Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            || | |   |   |   || |||   ||||||   || | ||   |     | ||  |||| |
Db    436 AAACAGGTTGTAGTGGAACACAAAGAAGCTTTGATCCAAGAAAACAGCATTGAAGAGGAT 495

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
          |||  || |     |  ||        | |    | |||| ||||| | |||  |        |  |
Db    496 GAACAAATGGATAAAGAAGAGAGTCCATTTAAAGGAATCGCTCCCATAAAGACTTCCAGC 555

Qy    643 CATATTATT 651
          | |||| ||
Db    556 CCTATTGTT 564


RESULT 27
AL716995
LOCUS       AL716995                 772 bp    mRNA    linear   EST 18-APR-2002
DEFINITION  AL716995 Danio rerio embryonic inner ear subtracted cDNA Danio
            rerio cDNA clone BN0AA021ZC10 3', mRNA sequence.
ACCESSION   AL716995
```

```
VERSION     AL716995.1  GI:20181598
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin
            ,J.P., Weissenbach,J. and Petit,C.
  TITLE     A subtracted cDNA library from the zebrafish (Danio rerio)
            embryonic inner ear
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..772
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="BN0AA021ZC10"
                     /clone_lib="Danio rerio embryonic inner ear subtracted
                     cDNA"
                     /tissue_type="inner ear"
                     /dev_stage="embryonic"
                     /note="subtracted cDNA library"
BASE COUNT      259 a    145 c    186 g    182 t
ORIGIN

  Query Match             1.3%;  Score 41.8;  DB 9;  Length 772;
  Best Local Similarity   51.3%;  Pred. No. 2;
  Matches   97;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy    463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
           | | | ||  | ||  ||| | | |     | ||    |  | |||| ||| | | |  ||
Db    332 GTAAATGGTGAAAAGAAAAGTGAAAGTTCATCTAGTTCAGAAATGCAGAATGCATCATCT 391

Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            || | |  |  |  || |||  ||||||  || | ||  |    | ||  |||| |
Db    392 AAACAGGTTGTAGTGGAACACAAAGAAGCTTTGATCCAAGAAAACAGCATTGAAGAGGAT 451

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
          |||  || |   |  ||     | |   | |||| ||||| | ||| |        |  |
Db    452 GAACAAATGGATAAAGAAGAGAGTCCATTTAAAGGAATCGCTCCCATAAAGACTTCCAGC 511

Qy    643 CATATTATT 651
           | |||| ||
Db    512 CCTATTGTT 520


RESULT 28
AL716744
LOCUS       AL716744                 790 bp    mRNA    linear   EST 18-APR-2002
DEFINITION  AL716744 Danio rerio embryonic inner ear subtracted cDNA Danio
            rerio cDNA clone BN0AA026ZG05 3', mRNA sequence.
ACCESSION   AL716744
VERSION     AL716744.1  GI:20181347
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin
            ,J.P., Weissenbach,J. and Petit,C.
  TITLE     A subtracted cDNA library from the zebrafish (Danio rerio)
            embryonic inner ear
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..790
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="BN0AA026ZG05"
                     /clone_lib="Danio rerio embryonic inner ear subtracted
                     cDNA"
                     /tissue_type="inner ear"
                     /dev_stage="embryonic"
                     /note="subtracted cDNA library"
BASE COUNT      264 a    153 c    192 g    181 t
ORIGIN

  Query Match             1.3%;  Score 41.8;  DB 9;  Length 790;
  Best Local Similarity   51.3%;  Pred. No. 2;
  Matches   97;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy    463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
           | | | ||  | ||  ||| | | |     | ||    |  | |||| ||| | | |  ||
Db    333 GTAAATGGTGAAAAGAAAAGTGAAAGTTCATCTAGTTCAGAAATGCAGAATGCATCATCT 392

Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            || | |  |  |  || |||  ||||||  || | ||  |    | ||  |||| |
Db    393 AAACAGGTTGTAGTGGAACACAAAGAAGCTTTGATCCAAGAAAACAGCATTGAAGAGGAT 452

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
          |||  || |   |  ||     | |   | |||| ||||| | ||| |        |  |
Db    453 GAACAAATGGATAAAGAAGAGAGTCCATTTAAAGGAATCGCTCCCATAAAGACTTCCAGC 512

Qy    643 CATATTATT 651
           | |||| ||
Db    513 CCTATTGTT 521


RESULT 29
AL717085/c
LOCUS       AL717085                 792 bp    mRNA    linear   EST 18-APR-2002
DEFINITION  AL717085 Danio rerio embryonic inner ear subtracted cDNA Danio
            rerio cDNA clone BN0AA021ZC10 5', mRNA sequence.
ACCESSION   AL717085
VERSION     AL717085.1  GI:20181688
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin
            ,J.P., Weissenbach,J. and Petit,C.
  TITLE     A subtracted cDNA library from the zebrafish (Danio rerio)
            embryonic inner ear
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..792
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="BN0AA021ZC10"
                     /clone_lib="Danio rerio embryonic inner ear subtracted
                     cDNA"
                     /tissue_type="inner ear"
                     /dev_stage="embryonic"
                     /note="subtracted cDNA library"
BASE COUNT      177 a    194 c    165 g    256 t
ORIGIN

  Query Match             1.3%;  Score 41.8;  DB 9;  Length 792;
  Best Local Similarity   51.3%;  Pred. No. 2;
```

```
  Matches    97;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy      463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
             | | | ||   | ||   ||| | | |         | ||    |   | |||| ||| | | |   ||
Db      755 GTAAATGGTGAAAAGAAAAGTGAAAGTTCATCTAGTTCAGAAATGCAGAATGCATCATCT 696

Qy      523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
              || |  |   |     || |||    ||||||    || |  ||   |     | ||   ||||  |
Db      695 AAACAGGTTGTAGTGGAACACAAAGAAGCTTTGATCCAAGAAAACAGCATTGAAGAGGAT 636

Qy      583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
            |||  || |     |   ||       | |     | ||||  |||||   | |||   |          |   |
Db      635 GAACAAATGGATAAAGAAGAGAGTCCATTTAAAGGAATCGCTCCCATAAAGACTTCCAGC 576

Qy      643 CATATTATT 651
             | |||| ||
Db      575 CCTATTGTT 567
```

RESULT 30
CNS0052P
LOCUS CNS0052P 844 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR11P16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL056652
VERSION AL056652.1 GI:4932342
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 844)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1. .844
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR11P16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      261 a     112 c      92 g      35 t     344 others
ORIGIN

  Query Match              1.3%;  Score 41.8;  DB 17;  Length 844;
  Best Local Similarity   16.4%;  Pred. No. 2;
  Matches   51;  Conservative  130;  Mismatches  130;  Indels    0;  Gaps    0;

Qy     2199 CCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGT 2258
            |:||| | : | ||:::::|   ::::::  :   :::   :|:  : :|   ::::||
Db      383 CBGCAAANMCACGCSSSSSASRSRMMSSCSSCASSSCSVGSSSCSAMSSSCCCVVSSCGM 442

Qy     2259 ACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTA 2318
            |: :|  |    | ::::: |:::: | : : :     |   ::   | |::: | :||: : :
Db      443 ASSSCCSGGMSASSSSSSGGVSSVSGGRAVGRGRCMCCSMVCCCMCSSMCCSMCCSVSSV 502

Qy     2319 TGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGA 2378
             :|:   :   : : |::   : |  ::: :  |   :|:   |::::: ::   :::: : |
Db      503 CAVCSSGSVRSVAVGCSVGGRVGGSSRGRAGRSSGRGGGSSVSSGVVSSSSSVGMGCACS 562

Qy     2379 CGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGG 2438
             : ::   : :::    :::   :::|     :| |        :::::     ::   :  :| :|: ::
Db      563 SASSVSCSBSSVASSVSSGBVSRGCGRCVGGGVGGGSRVSSCSSGSSSSSSSGCSGVSRS 622

Qy     2439 AACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCAT 2498
             :::|    |:::  |:|  :| |    |::||   |:|::    ::   :   :   : |:   |   :::
Db      623 RRSGSANGVRVGGSGRRGGGGGGRRGGANARADDAAABRABKARANSNADAWGAAVSSSV 682

Qy     2499 TGCCGCGAAAA 2509
              :::|  :   ::
Db      683 SSSBGARNGRR 693
```

RESULT 31
CNS012I3/c
LOCUS CNS012I3 645 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC BACN08C07 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL101589
VERSION AL101589.1 GI:5613200
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 645)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

```
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN08C07"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT       28 a      26 c      85 g      92 t     414 others
ORIGIN

  Query Match              1.3%;  Score 41.6;  DB 17;  Length 645;
  Best Local Similarity   11.0%;  Pred. No. 2.1;
  Matches   54;  Conservative  124;  Mismatches  314;  Indels    0;  Gaps    0;

Qy     1852 GGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCC 1911
             :     |      |:::          | :    |                 | ||     ||  |   : : ::::
Db      645 SNNNNANNNNNASSSNSNNNNASSNNNANANANNNTTTTTTTNTTTTTTTTSSSNSSSS 586

Qy     1912 AAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCT 1971
             || ::: ::    |        |           :          :       |    :::|    :       |    :     :
Db      585 SCAAASSSNSSNCAGSNNSCNGSNNNSSNSNCSSVACAGSSSSANCSSACSCSNSNNNSN 526

Qy     1972 TCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTC 2031
              ::          :: |          |    |    ::          | |::        ::        :: |    |    |
```

```
Db     525 NNSSNNNSNSSSATNNNNCTANAANASSANNTSAANGSSNSNNSSAAANSSNAASASNTT 466

Qy    2032 CGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTG 2091

Db     465 NSSSSNCNNNNSNNNSSTSSNAAAASSSSCASNSSSSSASASNAASSSSAASSANSNSSN 406

Qy    2092 AAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCC 2151

Db     405 SASSASANSSASSSSSSSSSSCSSCNASSAATSNSNNSNSAVSNNSTNNNNNNNSSSNSSS 346

Qy    2152 TCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATG 2211

Db     345 SSSSCSVAACCSSSNSSSSCATSNNSSASNSANAASASSSSNNSNNSNASSNNNNSSNNN 286

Qy    2212 CCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAAT 2271

Db     285 SNANNSNNNSSANSNSSSSSCSCNNSNSNNSSSCCSNNNNSSSCSSSNNNNSNNNSNNSN 226

Qy    2272 GCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACC 2331

Db     225 NNNNANNNSNSSSSSSNAANANNNNGNAGNSGGGSNNNNNCCNNNAAACCNANNAANNNC 166

Qy    2332 GCCGCCCATGCC 2343

Db     165 CNNNNNCCCCCC 154
```

RESULT 32
BQ759294

```
LOCUS       BQ759294                 444 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  EBpi01_SQ004_H04_R pistil, 1 DPA, no treatment, cv Optic, EBpi01
            Hordeum vulgare cDNA clone EBpi01_SQ004_H04 5', mRNA sequence.
ACCESSION   BQ759294
VERSION     BQ759294.1  GI:21967766
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
            Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  TITLE     Development of Barley Transcriptome Resources
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Waugh R, Marshall DF
            Genome Dynamics/Computational Biology
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: est@scri.sari.ac.uk.
FEATURES             Location/Qualifiers
     source          1..444
                     /organism="Hordeum vulgare"
                     /cultivar="Optic"
                     /db_xref="taxon:4513"
                     /clone="EBpi01_SQ004_H04"
                     /clone_lib="pistil, 1 DPA, no treatment, cv Optic, EBpi01"
                     /tissue_type="pistil"
                     /dev_stage="1 DPA"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                     Non-normalised library, directionally cloned into pSPORT1.
                     Derived from pistils dissected from developing grains (24
                     hours post anthesis) in glasshouse grown barley plants.
                     Developed as part of the barley transcriptome resources of
                     BBSRC/SEERAD funded cereal IGF (Investigating Gene
                     Function) project."
BASE COUNT       74 a    177 c    137 g     56 t
ORIGIN

  Query Match             1.3%;  Score 41.4;  DB 14;  Length 444;
  Best Local Similarity   55.1%;  Pred. No. 2;
  Matches   81;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy    1219 ATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAG 1278

Db     235 ATCTCCGGGACGTCCATGTCGTGCCCGCACGTGAGCGGCCTCGCCGCGCTGCTCCGGCAG 294

Qy    1279 AAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGAC 1338

Db     295 GCGCACCCGGACTGGAGCCCCGCGGAGGTCAAGTCGGCGCTCATGACCACGGCGTACAAC 354

Qy    1339 ATCGGTGCAGTCGGCGTGGACAGCAAG 1365

Db     355 CTGGACAACTCCGGGGAGATCATCAAG 381
```

RESULT 33
BE778605

```
LOCUS       BE778605                 976 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601466144F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3869242 5',
            mRNA sequence.
ACCESSION   BE778605
VERSION     BE778605.1  GI:10199803
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 976)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9618  row: h  column: 11
            High quality sequence start: 7
            High quality sequence stop: 561.
FEATURES             Location/Qualifiers
     source          1..976
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3869242"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      259 a    331 c    276 g    110 t
ORIGIN

  Query Match             1.3%;  Score 41.4;  DB 12;  Length 976;
  Best Local Similarity   50.8%;  Pred. No. 2.9;
  Matches   99;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy    2375 CGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACG 2434

Db     730 CGGACCGGCTAGCCCGGCTCGCCACAGGCCTACGCCCGACGGCCCTCAGGACCCAGGCGG 789

Qy    2435 GTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCG 2494

Db     790 ACGGACCGTGTGAACATAGCAGCGCCCGACCAACCACAGGCACGAACACACAGCCGCACC 849

Qy    2495 GCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACA 2554
```

```
Db     850 CCAAGCGAGCCACACACACGACACCAACGAAACAAACCAACACAAAGCGCGCGAAACACA 909

Qy    2555 GCACATGGAGCGAAA 2569
             ||||  | ||  ||
Db     910 ACACACAGCGCCCAA 924


RESULT 34
BI778455
LOCUS       BI778455                 588 bp    mRNA    linear   EST 23-JUL-2002
DEFINITION  EBro07_SQ003_F16_R root, 3 week, reduced light, cv Optic, EBro07
            Hordeum vulgare cDNA clone EBro07_SQ003_F16 5', mRNA sequence.
ACCESSION   BI778455
VERSION     BI778455.2  GI:21950134
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
            Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  TITLE     Development of Barley Transcriptome Resources
  JOURNAL   Unpublished (2001)
COMMENT     On Sep 26, 2001 this sequence version replaced gi:15781347.
            Contact: Waugh R, Marshall DF
            Genome Dynamics/Computational Biology
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: est@scri.sari.ac.uk
            All sequence has a Phred quality score of 20 or over
            Seq primer: M13 reverse.
FEATURES             Location/Qualifiers
     source          1..588
                     /organism="Hordeum vulgare"
                     /cultivar="Optic"
                     /db_xref="taxon:4513"
                     /clone="EBro07_SQ003_F16"
                     /clone_lib="root, 3 week, reduced light, cv Optic, EBro07"
                     /tissue_type="root"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                     Non-normalised library, directionally cloned into pSPORT1.
                     Derived from shoot and root material of 3 week old
                     etiolated barley plants.  Developed as part of the barley
                     transcriptome resources of BBSRC/SEERAD funded cereal IGF
                     (Investigating Gene Function) project."
BASE COUNT      102 a    205 c    199 g     82 t
ORIGIN

  Query Match             1.3%;  Score 41.2;  DB 13;  Length 588;
  Best Local Similarity   56.7%;  Pred. No. 2.6;
  Matches   76;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy    1219 ATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAG 1278
           ||  |||| |||||| | ||      ||    ||| | ||||    ||||| ||||| | |
Db     454 ATCTCCGGCACATCCATGTCGTGCCCGCACGTGAGCGGCCTCGCGGCGCTGCTCCGGAGC 513

Qy    1279 AAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGAC 1338
                 |||||  |  ||||  || |  | ||||  |  ||||  |||| || ||  |    |
Db     514 GCGCACCCGGAGTGGAGCCCCGCCGCCGTGCGCTCGGCGCTCATGACGACCGCCTACTCC 573

Qy    1339 ATCGGTGCAGTCGG 1352
           |     | | |||
Db     574 ACGTACACCGGCGG 587


RESULT 35
CNS021SP
LOCUS       CNS021SP                 847 bp    DNA     linear   GSS 12-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            226G16 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL177298
VERSION     AL177298.1  GI:7815355
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 847)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 847)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 847)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..847
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="226G16"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : COAG226BD08LP1-end : T7"
BASE COUNT      199 a    209 c    231 g    203 t      5 others
ORIGIN

  Query Match             1.3%;  Score 41.2;  DB 17;  Length 847;
  Best Local Similarity   49.1%;  Pred. No. 3;
  Matches  106;  Conservative    1;  Mismatches  109;  Indels    0;  Gaps    0;

Qy    2635 ATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGC 2694
           ||| ||      ||  |||    | | ||| |      ||   |||  | || || || | ||| |
Db     182 ATCAGCATCGGCATCGGCACCTGCACCTGCACCTGCAGCAACAGCATCAGCACCGGCACC 241

Qy    2695 ACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCA 2754
              | |       |     |||  |     ||| |||      || | |  || |    |  | |
Db     242 TGCAGCATTAGCCGCGATGATGCTAGCATCGTAGGGGGAGGAGGAGGATGGGGCGACAG 301

Qy    2755 CTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTG 2814
           |||  |    |  ||   || | | | ||    | || |   | || : ||||| | ||||
Db     302 CTGCACCCCGCAAATTTTGCATCTCCCAGCAGGGCATTTATGCMCGTCGACGCCGGTGGA 361

Qy    2815 CGCTACGACCTGCTCAAACAGGATGCATTCGCCGAA 2850
              ||   ||   |   |    | || |||||||    || |
Db     362 TGCGGGGATTTCGTACGAGAGTATGCATTTCACGCA 397


RESULT 36
BG368419
LOCUS       BG368419                 573 bp    mRNA    linear   EST 22-OCT-2001
DEFINITION  HVSMEi0018I17f Hordeum vulgare 20 DAP spike EST library HVcDNA0010
            (20 DAP) Hordeum vulgare cDNA clone HVSMEi0018I17f, mRNA sequence.
ACCESSION   BG368419
VERSION     BG368419.1  GI:13257520
```

```
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
            ,R.D., Close,S.J., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Morex 20 DAP spike cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 238
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 309.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Hordeum vulgare"
                     /cultivar="Morex"
                     /db_xref="taxon:4513"
                     /clone="HVSMEi0018I17f"
                     /clone_lib="Hordeum vulgare 20 DAP spike EST library
                     HVcDNA0010 (20 DAP)"
                     /tissue_type="20 DAP spike"
                     /lab_host="SOLR"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
                     Plants were grown in the greenhouse at the University of
                     California, Riverside (Fenton, SJ Close, TJ Close). Whole
                     spikes with awns trimmed were collected at 20 DAP (Fenton
                     ). Total RNA was prepared, poly(A) RNA was purified, one
                     primary unamplified cDNA library was made, and 1 million
                     pfu were in vivo excised to give pBluescript SK(-) cDNA
                     phagemids in the TJ Close lab at the University of
                     California, Riverside (Choi). Phagemids were plated and
                     picked at the Clemson University Genomics Institute (CUGI)
                     (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
                     preparations, DNA sequencing and sequence analysis were
                     performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
                     , Rambo, Main). The sequence has been trimmed to remove
                     vector sequence and contains a minimum of 100 bases of
                     phred value 20 or above. For more details on library
                     preparation and sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT      126 a    176 c    193 g     78 t
ORIGIN

  Query Match             1.3%;  Score 41;  DB 12;  Length 573;
  Best Local Similarity   50.8%;  Pred. No. 2.9;
  Matches   98;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;

Qy     2973 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 3032
             ||| | |  || | |||| | || || | | || ||||||| | |  ||| |  | |
Db      126 GGAGCCCATCCAGTACGGGCACGTCTTCGCCGTCACGGGCGACCTGGCCGTCCAGGCCAT 185

Qy     3033 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGG 3092
             ||  |  |  |  ||| |||         |   |  |      || ||   | ||  | ||
Db      186 CGCGCCACGGGACGCGGAGGCCATGCGCACCGCCGAGGAGAGCGTGCAGGGCGTGCAGGT 245

Qy     3093 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 3152
            | || | | | |   ||||  |  | |   || | |||        | |     ||||| |
Db      246 CCCGCAGGCCAACGGCGGCGGCTTCAGCGCCGCCGTGGGCATGGAGACGGTGGCCGGGTA 305

Qy     3153 CAAACAGTACGGC 3165
            ||| |||  ||||
Db      306 CAACCAGGCCGGC 318


RESULT 37
AW006888
LOCUS       AW006888                 540 bp    mRNA    linear   EST 08-MAR-2000
DEFINITION  ws15f06.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2497283 3'
            similar to SW:TISD_HUMAN P47974 TIS11D PROTEIN ;contains MER22.t3
            TAR1 repetitive element ;, mRNA sequence.
ACCESSION   AW006888
VERSION     AW006888.1  GI:5855666
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 540)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 747   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 479.
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2497283"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT       79 a    212 c    177 g     72 t
ORIGIN

  Query Match             1.3%;  Score 40.8;  DB 10;  Length 540;
  Best Local Similarity   47.1%;  Pred. No. 3.2;
  Matches  161;  Conservative    0;  Mismatches  177;  Indels    4;  Gaps    1;

Qy     2152 TCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATG 2211
            || |  | || | ||| || | | | | |   |     || ||||  | || |  |    |
Db      195 TCAGCAGCAGCAGCAGCAGCAGGGCCTGGCGCCCCCGCGCAGCCGCCGGCGCCGCCCAG 254

Qy     2212 CCGGGCATCCGCCCC----TACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
            | |  |   | || |       |||||||||  | | | ||   |||||||  |  |  |
Db      255 CGCGACCCTCCCCGCCGGGGCCGGCGCAACTCCCTGAGCTCCGGCAGCCTCAGCGGCTCT 314

Qy     2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
            || |  | | | |     |     |  |  | | |   |   |  ||| || | |
Db      315 GAGTCTCCCAGCCTCGACCCTGGCCGCCGCCTGCCAATCTTCAGCCGCCTCTCCATCTCC 374
```

```
Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
          || |  |        |        ||| || |   |    | ||| ||  || ||    ||||||
Db    375 GACGACTGAGGCAAGAGGGCGCCAGTGAGGAGGAAGGGAAGGCGGTTCAGAGATGTTGGA 434

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
             ||| | ||     |||   | |     | |   ||    |  ||   || |   || ||||
Db    435 GGACACCCCTCGCCATCTCGCCCTTGCTGGGGGCACGGGAGTGGGGGGGTGACATGGGC 494

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAAC 2489
            |  |||  |  |  ||| |   |   | ||     | || | |||
Db    495 CCTAGGCAGACTGCAAGCCCGACCGAGCACTTGGACTCGAAC 536
```

RESULT 38
BI790356/c
LOCUS   BI790356   631 bp   mRNA   linear   EST 12-MAR-2002
DEFINITION id06e08.x1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus
            musculus cDNA clone IMAGE:5662550 3' similar to TR:Q13151 Q13151
            HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A0 ;, mRNA sequence.
ACCESSION   BI790356
VERSION   BI790356.1 GI:15818081
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 631)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE   Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu)
            MGI:1948876 This sequence now available from the IMAGE consortium,
            for clone orders contact: info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 462.
FEATURES   Location/Qualifiers
     source   1. .631
                     /organism="Mus musculus"
                     /strain="ICR"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5662550"
                     /clone_lib="Melton Normalized Mixed Mouse Pancreas 1
                     N1-MMS1"
                     /sex="Both for embryonic & newborn, male for adult and
                     adult islet"
                     /dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
                     adult, mixed"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; Five
                     libraries representing EI0.5/12.5 pancreatic bud, E16.5
                     pancreas, newborn pancreas, adult pancreas, and adult
                     islets of Langerhans were seperately constructed using
                     SuperScript Plasmid Library kit (Life Technologies). cDNA
                     was made by oligo-dT priming and size-selected by column
                     fractionation. Libraries were amplified once on solid
                     support and plasmid DNA from each library was prepared
                     and mixed in equal amounts. The mixed library DNA was
                     normalized by method #4 from Bonaldo, Lennon, and Soares
                     1996 Genome Research 6:791-806; 0.5 microgram
                     single-stranded mixed library plasmid DNA was mixed with
                     5 micrograms PCR product representing mixed library
                     inserts and hybridized to an Ecot of 6. Single-stranded
                     (unhybridized) plasmids were isolated by hydroxyapatite
                     chromatography and used to make this library."
BASE COUNT   76 a   267 c   185 g   102 t   1 others
ORIGIN

  Query Match   1.3%; Score 40.8; DB 13; Length 631;
  Best Local Similarity   48.3%; Pred. No. 3.5;
  Matches 114; Conservative   0; Mismatches 122; Indels   0; Gaps   0;

```
Qy   2968 GGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCG 3027
           ||||  |||||  ||||  ||||||  |   |  |  |       |        ||||||       ||||
Db    281 GGCGGCGAACGAGACCAGAACGGGCTGGCCAAGGGCGGCGGTGGCGGCGGCGGCGGCTAC 222

Qy   3028 ACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGT 3087
          |            |||| |    |||        ||             |  |       ||    |   | |||
Db    221 AACAGCTACGGCGGCTACGGGGGCTACGGCGCCTACGGGGGCGGTGGCGGCGGCGGCGGC 162

Qy   3088 CTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCC 3147
                 ||  |       ||| | ||||||||||||        |||||| |     | |||||| |    |
Db    161 TCATACGGCGGCAGCGACTACGGCAACGGCTTCGGCGGCTTCGGCAGCTACAGCCAGCAC 102

Qy   3148 GGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTG 3203
              |||       |            |      |  |||||  |||    ||||    |||   | |    |||
Db    101 CAGTCCTCGTACGGGCCGATGAAGAGCGGCGGAGGAGGCGGCGGCGGCGGCAGCTG 46
```

RESULT 39
BM068591
LOCUS   BM068591   474 bp   mRNA   linear   EST 13-NOV-2001
DEFINITION WHE3459_D08_G15ZS Wheat pre-anthesis spike cDNA library Triticum
            aestivum cDNA clone WHE3459_D08_G15, mRNA sequence.
ACCESSION   BM068591
VERSION   BM068591.1 GI:16907769
KEYWORDS   EST.
SOURCE   bread wheat.
  ORGANISM Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 474)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE   The structure and function of the expressed portion of the wheat
            genomes - Pre-anthesis spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES   Location/Qualifiers
     source   1. .474
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE3459_D08_G15"
                     /clone_lib="Wheat pre-anthesis spike cDNA library"
                     /tissue_type="Spike before anthesis"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;

```
                Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                greenhouse.  Whole spike with awns trimmed,  white, green
                and yellow anther were collected and total RNA, and
                poly(A) RNA were prepared, a cDNA library was made, and
                the cDNA  clones were in vivo excised to give pBluescript
                phagemids in the TJ  Close lab (Choi, Close, Fenton) at
                the University of California,  Riverside.  Plasmid DNA
                preparations and DNA sequencing were performed  in the OD
                Anderson lab (all other authors)."
BASE COUNT      182 a    168 c     49 g     75 t
ORIGIN

  Query Match             1.3%;  Score 40.4;  DB 13;  Length 474;
  Best Local Similarity   51.7%;  Pred. No. 4;
  Matches   92;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy        8 CGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCTGTTGCAA 67
              | ||| ||||| |  || | | ||||  || | | || ||||   |   |   |     |||
Db      254 CAACCGCAACCATTTCTGCCACAACTACCATATCCGCAGCCACAACCATTTCCACCACAA 313

Qy       68 CAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAG 127
            ||||  | ||  |    |   |   |    ||  ||   |   |  || | ||||  |||
Db      314 CAACCATATCCACAACCGCAACCACAGTATCCGCAACCACAACAACCAATTTCACAGCAA 373

Qy      128 GCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATC 185
                ||||   | |   |  |  |  | | |||||||| | | | || || ||  | |
Db      374 CAAGCACAACAAGCACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 431


RESULT 40
AG058272
LOCUS       AG058272                1600 bp    DNA     linear   GSS 02-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-045B19.R, genomic survey sequence.
ACCESSION   AG058272
VERSION     AG058272.1  GI:16595733
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
            BAC Library clone:PTB-045B19.R.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1600)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1..1600
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="PTB-045B19.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      332 a    417 c    525 g    298 t     28 others
ORIGIN

  Query Match             1.3%;  Score 40.4;  DB 17;  Length 1600;
  Best Local Similarity   48.3%;  Pred. No. 7;
  Matches  101;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy     2231 GCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCA 2290
             || | |     || ||||| |          |||  |        || | ||  | | |  |
Db     1108 GCCCCATCNCGCGAGCAGCTGTGTGGTATCAGATTAGATTGTCCACGGATCGCGCAGTCG 1167

Qy     2291 TCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGC 2350
               || |      ||||||  |||    |   || ||| ||||  | |      | | ||
Db     1168 CGGTCGAGTCGGCCGCCGCGCGCGTAAGNGATTACGGTANCGCCAACGANTATCACAGGC 1227

Qy     2351 AGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCG 2410
            |||  ||| | | ||| ||||  ||| |  |||  |         || ||| | |    |
Db     1228 AGGCNCGCAGACNGAATGCCGGCTCGCAACGGTATGTGTCTGGAGCGACGCGACGCAGAC 1287

Qy     2411 TCATCGCGCAAACCCAACAGGACGGTGGA 2439
              || |||||   | | ||   |   |||| |
Db     1288 GCAGCGCGCGCCCGCGACGCGNTGGTGTA 1316

Search completed: January 27, 2003, 19:09:14
Job time : 2951 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 20:44:39 ; Search time 5463 Seconds
(without alignments)
17068.532 Million cell updates/sec

Title: US-09-830-433A-7
Perfect score: 3204
Sequence: 1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3204	100.0	3204	6	AX024064	AX024064 Sequence
	2	3153	98.4	3204	1	NME311654	AJ311654 Neisseria
c	3	2156	67.3	349061	1	NMA2Z2491	AL162753 Neisseria
	4	834	26.0	4218	6	AX236469	AX236469 Sequence
	5	834	26.0	4218	6	AX239755	AX239755 Sequence
	6	834	26.0	10591	1	AE002545	AE002545 Neisseria
	7	834	26.0	349980	6	AX044034	AX044034 Sequence
	8	831	25.9	3939	6	AX236409	AX236409 Sequence
	9	831	25.9	3939	6	AX236419	AX236419 Sequence
	10	831	25.9	3939	6	AX239717	AX239717 Sequence
	11	831	25.9	3939	6	AX239729	AX239729 Sequence
	12	831	25.9	4170	6	AX236463	AX236463 Sequence
	13	831	25.9	4170	6	AX239749	AX239749 Sequence
	14	831	25.9	4179	6	AX236413	AX236413 Sequence
	15	831	25.9	4179	6	AX239721	AX239721 Sequence
	16	831	25.9	4335	6	AX236457	AX236457 Sequence
	17	831	25.9	4335	6	AX239743	AX239743 Sequence
	18	831	25.9	4344	6	AX236411	AX236411 Sequence
	19	831	25.9	4344	6	AX239719	AX239719 Sequence
	20	831	25.9	4425	6	AX236407	AX236407 Sequence
	21	831	25.9	4425	6	AX239715	AX239715 Sequence
	22	783	24.4	3254	1	NME277537	AJ277537 Neisseria
c	23	321	10.0	600	1	AF169448	AF169448 Neisseria
	24	177	5.5	286	1	AF169473	AF169473 Neisseria
	25	177	5.5	286	6	A68924	A68924 Sequence 95
	26	35	1.1	50	6	AX236815	AX236815 Sequence
	27	33	1.0	1434	1	NGOPAK	X52364 N.gonorrhoe
c	28	25	0.8	37497	2	AC017356	AC017356 Drosophil
c	29	25	0.8	161278	3	AC009749	AC009749 Drosophil
	30	25	0.8	327209	3	AE003575	AE003575 Drosophil
c	31	23	0.7	30	6	AX024142	AX024142 Sequence
c	32	23	0.7	14043	1	AB014075	AB014075 Clostridi
c	33	22	0.7	48626	2	AC013119	AC013119 Drosophil
c	34	22	0.7	85918	2	DMBR33H2	AL121812 Drosophil
	35	22	0.7	160986	3	AC108481	AC108481 Drosophil
	36	22	0.7	267113	2	AC099761	AC099761 Mus muscu
	37	22	0.7	294218	3	AE003430	AE003430 Drosophil
	38	21	0.7	34	6	AX024141	AX024141 Sequence
c	39	21	0.7	348	5	AF281339	AF281339 Oncorhync
c	40	21	0.7	768	9	HSA332365	AJ332365 Homo sapi
	41	21	0.7	12211	1	AE004895	AE004895 Pseudomon
c	42	21	0.7	121118	2	AC114017	AC114017 Rattus no
c	43	21	0.7	124667	2	AC015568	AC015568 Homo sapi
c	44	20	0.6	735	6	AR140161	AR140161 Sequence
c	45	20	0.6	735	6	AR199650	AR199650 Sequence
c	46	20	0.6	735	6	E16589	E16589 Serratia fi
c	47	20	0.6	5019	6	AX106914	AX106914 Sequence
	48	20	0.6	10034	1	AE002016	AE002016 Deinococc
c	49	20	0.6	12250	1	AE007837	AE007837 Clostridi
	50	20	0.6	20409	2	AC014557	AC014557 Drosophil
	51	20	0.6	36328	3	LMFL5213	AL583932 Leishmani
c	52	20	0.6	89047	6	AX067459	AX067459 Sequence
c	53	20	0.6	118567	2	AC114134	AC114134 Rattus no
	54	20	0.6	127506	8	OSJN00096	AL606647 Oryza sat
c	55	20	0.6	150962	2	AL627100	AL627100 Homo sapi
	56	20	0.6	152763	2	AC119715	AC119715 Rattus no
	57	20	0.6	152985	9	AC020581	AC020581 Homo sapi
	58	20	0.6	157448	2	AC131139	AC131139 Rattus no
	59	20	0.6	159717	2	AC008368	AC008368 Trypanoso
c	60	20	0.6	159717	2	AC008368	AC008368 Trypanoso
c	61	20	0.6	160551	9	AC009329	AC009329 Human Chr
c	62	20	0.6	161252	2	AC127258	AC127258 Mus muscu
c	63	20	0.6	163115	2	AC101752	AC101752 Mus muscu
c	64	20	0.6	164942	3	AC011758	AC011758 Drosophil
	65	20	0.6	173722	2	AL450993	AL450993 Homo sapi

c	66	20	0.6	175621	2	AC053486	AC053486 Homo sapi
c	67	20	0.6	177182	10	AL669946	AL669946 Mouse DNA
	68	20	0.6	180247	2	AC053509	AC053509 Homo sapi
c	69	20	0.6	185447	9	AL450992	AL450992 Human DNA
	70	20	0.6	191053	10	AL672249	AL672249 Mouse DNA
	71	20	0.6	204702	9	AL356957	AL356957 Human DNA
	72	20	0.6	216686	2	AC122873	AC122873 Mus muscu
c	73	20	0.6	302303	3	AE003571	AE003571 Drosophil
	74	19	0.6	29	6	AX236809	AX236809 Sequence
	75	19	0.6	32	6	AX236813	AX236813 Sequence
c	76	19	0.6	126	6	AX017688	AX017688 Sequence
c	77	19	0.6	300	11	G13553	G13553 human STS S
c	78	19	0.6	426	4	AY011990	AY011990 Sorex ara
	79	19	0.6	654	9	HSM802875	AL512738 Homo sapi
	80	19	0.6	807	6	AX119139	AX119139 Sequence
	81	19	0.6	1061	8	AF053077	AF053077 Nicotiana
c	82	19	0.6	1265	9	HUMPROTZ	M59303 Human vitam
	83	19	0.6	1286	6	BD012333	BD012333 Genes rel
	84	19	0.6	1286	9	AK027731	AK027731 Homo sapi
	85	19	0.6	1286	23	BD005013	Bd005013 Genes rel
	86	19	0.6	1296	9	BC012379	BC012379 Homo sapi
	87	19	0.6	1383	6	AX119157	AX119157 Sequence
	88	19	0.6	1451	9	AK000962	AK000962 Homo sapi
c	89	19	0.6	1485	9	HUMPROZI	M55670 Human prote
c	90	19	0.6	1551	6	AX409708	AX409708 Sequence
c	91	19	0.6	1551	9	HUMPROZII	M55671 Human prote
	92	19	0.6	1596	6	AX119159	AX119159 Sequence
c	93	19	0.6	1700	10	BC027427	BC027427 Mus muscu
	94	19	0.6	1781	6	AX400011	AX400011 Sequence
	95	19	0.6	1828	9	AK000382	AK000382 Homo sapi
	96	19	0.6	1866	9	BC001246	BC001246 Homo sapi
c	97	19	0.6	2342	9	AB033749	AB033749 Homo sapi
	98	19	0.6	2796	3	AF043944	AF043944 Mytilus e
	99	19	0.6	3050	9	BC024153	BC024153 Homo sapi
	100	19	0.6	3127	10	BC004674	BC004674 Mus muscu
	101	19	0.6	3244	9	D50912	D50912 Human mRNA
	102	19	0.6	3301	9	HSM802123	AL137421 Homo sapi
	103	19	0.6	3402	9	AK055116	AK055116 Homo sapi
	104	19	0.6	3418	9	BC004181	BC004181 Homo sapi
	105	19	0.6	3423	9	BC008733	BC008733 Homo sapi
c	106	19	0.6	3889	8	PSY250467	AJ250467 Pinus syl
	107	19	0.6	4195	9	AB058700	AB058700 Homo sapi
c	108	19	0.6	7902	9	HUMDNAAPK	L27425 Homo sapien
	109	19	0.6	8902	2	AC017749	AC017749 Drosophil
	110	19	0.6	10029	1	AE013261	AE013261 Methanosa
c	111	19	0.6	10670	1	AE009117	AE009117 Agrobacte
c	112	19	0.6	10695	1	AE007265	AE007265 Sinorhizo
	113	19	0.6	10871	1	AE004643	AE004643 Pseudomon
c	114	19	0.6	11516	1	AE008082	AE008082 Agrobacte
	115	19	0.6	11627	1	AE012368	AE012368 Xanthomon
c	116	19	0.6	16418	9	AF440358	AF440358 Homo sapi
c	117	19	0.6	19482	1	AE007072	AE007072 Mycobacte
c	118	19	0.6	30000	6	AX250261	AX250261 Sequence
c	119	19	0.6	34150	1	MTCY190	Z70283 Mycobacteri
c	120	19	0.6	36348	3	AC024765	AC024765 Caenorhab
c	121	19	0.6	37918	9	AP005232	AP005232 Homo sapi
	122	19	0.6	38080	9	HS429E7	AL031722 Human DNA
	123	19	0.6	38141	2	AC018202	AC018202 Drosophil
c	124	19	0.6	40327	9	AP005379	AP005379 Homo sapi
	125	19	0.6	45998	9	HSU81031	U81031 Homo sapien
	126	19	0.6	62909	6	AX067457	AX067457 Sequence
c	127	19	0.6	69368	2	AC118642	AC118642 Mus muscu
	128	19	0.6	71994	2	AC123649	AC123649 Mus muscu
c	129	19	0.6	72179	2	AC096066	AC096066 Rattus no
	130	19	0.6	82849	3	AC004277	AC004277 Drosophil
c	131	19	0.6	86896	1	RCU57682	U57682 Rhodobacter
	132	19	0.6	89453	8	AC068655	AC068655 Genomic S
c	133	19	0.6	93896	8	AB046435	AB046435 Arabidops
c	134	19	0.6	93896	8	F15I15	AF296826 Arabidops
c	135	19	0.6	104170	3	AC091510	AC091510 Leishmani
	136	19	0.6	106763	9	AP002091	AP002091 Homo sapi
c	137	19	0.6	110000	2	LMFLCHR36_13	Continuation (14 o
	138	19	0.6	118860	2	AP004084	AP004084 Oryza sat
	139	19	0.6	122122	2	AC121621	AC121621 Rattus no
	140	19	0.6	124636	2	AC091690	AC091690 Oryza sat
	141	19	0.6	126169	9	AC010291	AC010291 Homo sapi
	142	19	0.6	127982	2	AC104473	AC104473 Oryza sat
c	143	19	0.6	132933	9	AL137002	AL137002 Human DNA
	144	19	0.6	134764	9	AP003305	AP003305 Homo sapi
c	145	19	0.6	139287	2	AC060794	AC060794 Homo sapi
c	146	19	0.6	141073	2	AP005512	AP005512 Oryza sat
	147	19	0.6	145456	9	AL513366	AL513366 Human DNA
c	148	19	0.6	145882	2	AL139131	AL139131 Homo sapi
	149	19	0.6	146927	2	AC020144	AC020144 Drosophil
	150	19	0.6	150643	2	AC102184	AC102184 Mus muscu
	151	19	0.6	151024	2	AC016855	AC016855 Homo sapi
	152	19	0.6	151031	9	AC112222	AC112222 Homo sapi
	153	19	0.6	151712	9	AC018992	AC018992 Homo sapi
	154	19	0.6	152560	9	AC011339	AC011339 Homo sapi
	155	19	0.6	153388	8	AP003289	AP003289 Oryza sat
	156	19	0.6	153557	2	AC019260	AC019260 Homo sapi
c	157	19	0.6	155383	2	AP004396	AP004396 Oryza sat
c	158	19	0.6	156551	2	AP003490	AP003490 Oryza sat
c	159	19	0.6	157722	3	AC092248	AC092248 Drosophil
	160	19	0.6	158932	2	AC126845	AC126845 Rattus no
	161	19	0.6	159101	2	AC126205	AC126205 Rattus no
c	162	19	0.6	159311	2	HS279F22	AL662879 Homo sapi
c	163	19	0.6	159734	2	AC118428	AC118428 Rattus no
	164	19	0.6	162010	2	AP004121	AP004121 Oryza sat
	165	19	0.6	163156	2	AP004085	AP004085 Oryza sat
	166	19	0.6	164970	2	AC098136	AC098136 Rattus no
	167	19	0.6	166512	3	AC099010	AC099010 Drosophil
	168	19	0.6	166727	2	AC130190	AC130190 Pan trogl
c	169	19	0.6	167115	9	AC090093	AC090093 Homo sapi
c	170	19	0.6	168058	3	AC010049	AC010049 Drosophil
	171	19	0.6	168267	8	AP003535	AP003535 Oryza sat
c	172	19	0.6	169414	9	AC068707	AC068707 Homo sapi
	173	19	0.6	169989	9	AC090137	AC090137 Homo sapi
	174	19	0.6	170232	9	AL135903	AL135903 Human DNA
	175	19	0.6	170330	9	AC104430	AC104430 Homo sapi
	176	19	0.6	173184	9	AC112515	AC112515 Homo sapi
c	177	19	0.6	174345	2	AC025236	AC025236 Homo sapi
	178	19	0.6	174646	2	AC087355	AC087355 Homo sapi
c	179	19	0.6	174974	3	AC092228	AC092228 Drosophil
	180	19	0.6	176123	9	AL390718	AL390718 Human DNA
	181	19	0.6	176562	2	AC103686	AC103686 Homo sapi
	182	19	0.6	178248	9	AC012476	AC012476 Homo sapi
c	183	19	0.6	180905	2	AC011688	AC011688 Homo sapi
c	184	19	0.6	181004	2	AC116431	AC116431 Papio cyn
	185	19	0.6	181727	2	AC036232	AC036232 Homo sapi
	186	19	0.6	183242	9	AC007906	AC007906 Homo sapi
c	187	19	0.6	183970	2	AC017022	AC017022 Homo sapi
c	188	19	0.6	184852	2	AC026032	AC026032 Homo sapi
	189	19	0.6	187023	2	AC115422	AC115422 Rattus no
c	190	19	0.6	187023	2	AC115422	AC115422 Rattus no
	191	19	0.6	187533	2	AP004806	AP004806 Oryza sat
	192	19	0.6	187658	2	AC098041	AC098041 Rattus no
	193	19	0.6	189055	2	AC073716	AC073716 Mus muscu
	194	19	0.6	191729	2	AL590872	AL590872 Homo sapi
c	195	19	0.6	192001	2	AC019139	AC019139 Homo sapi
c	196	19	0.6	193452	2	AC127575	AC127575 Mus muscu
	197	19	0.6	193828	9	AP003025	AP003025 Homo sapi
c	198	19	0.6	195640	2	AC094441	AC094441 Rattus no
c	199	19	0.6	196413	2	AC097269	AC097269 Pan trogl
	200	19	0.6	196668	2	AC117584	AC117584 Mus muscu
	201	19	0.6	197063	2	AL807240	AL807240 Mus muscu
c	202	19	0.6	200622	9	AC068228	AC068228 Homo sapi
	203	19	0.6	200799	8	AF137379	AF137379 Nephrosel
c	204	19	0.6	200799	8	AF137379	AF137379 Nephrosel
c	205	19	0.6	200990	2	AC099134	AC099134 Rattus no
c	206	19	0.6	202457	2	AC121759	AC121759 Homo sapi
c	207	19	0.6	203732	2	AC129320	AC129320 Mus muscu
c	208	19	0.6	204167	2	AC110553	AC110553 Mus muscu
	209	19	0.6	205050	1	AL646082	AL646082 Ralstonia
c	210	19	0.6	211544	9	AC025165	AC025165 Homo sapi
c	211	19	0.6	215188	2	AC124196	AC124196 Mus muscu

	212	19	0.6	216812	10	AL672073	AL672073 Mouse DNA
c	213	19	0.6	218277	2	AL671741	AL671741 Mus muscu
	214	19	0.6	222845	2	AC094653	AC094653 Rattus no
	215	19	0.6	227117	2	AC122341	AC122341 Mus muscu
	216	19	0.6	227391	2	AC125160	AC125160 Mus muscu
	217	19	0.6	237712	9	AC012634	AC012634 Homo sapi
	218	19	0.6	237952	2	AL591116	AL591116 Homo sapi
	219	19	0.6	238741	2	AC068894	AC068894 Homo sapi
c	220	19	0.6	240131	2	AL772361	AL772361 Mus muscu
c	221	19	0.6	240197	3	AE003643	AE003643 Drosophil
	222	19	0.6	242224	2	AC103337	AC103337 Rattus no
c	223	19	0.6	251997	2	AC012180	AC012180 Homo sapi
c	224	19	0.6	269823	3	AE003615	AE003615 Drosophil
	225	19	0.6	270150	9	AE006639	AE006639 Homo sapi
	226	19	0.6	272545	2	AC090533	AC090533 Mus muscu
c	227	19	0.6	305502	3	DROSADH03	AE003409 Drosophil
c	228	19	0.6	305516	3	AE003594	AE003594 Drosophil
	229	19	0.6	309805	2	AC026340	AC026340 Homo sapi
	230	19	0.6	338579	1	AP003004	AP003004 Mesorhizo
c	231	18	0.6	28	6	AX236792	AX236792 Sequence
c	232	18	0.6	28	6	AX239766	AX239766 Sequence
c	233	18	0.6	60	6	AR125920	AR125920 Sequence
c	234	18	0.6	60	6	I24287	I24287 Sequence 74
	235	18	0.6	111	6	AX079073	AX079073 Sequence
c	236	18	0.6	150	6	A01406	A01406 Malaria par
c	237	18	0.6	150	6	A13149	A13149 Nucleotide
c	238	18	0.6	150	6	A16111	A16111 oligonucleo
	239	18	0.6	213	6	AX436315	AX436315 Sequence
	240	18	0.6	224	8	AY022882	AY022882 Oryza sat
c	241	18	0.6	224	8	AY023142	AY023142 Oryza sat
	242	18	0.6	231	3	AF272155	AF272155 Procambar
c	243	18	0.6	328	11	G41605	G41605 Z15424 Zebr
	244	18	0.6	558	8	AB036884	AB036884 Arabidops
	245	18	0.6	588	6	AX381056	AX381056 Sequence
	246	18	0.6	661	6	AX381052	AX381052 Sequence
	247	18	0.6	687	9	HSA334262	AJ334262 Homo sapi
c	248	18	0.6	716	9	HSA338190	AJ338190 Homo sapi
	249	18	0.6	740	9	HSA336538	AJ336538 Homo sapi
	250	18	0.6	772	9	HSA341989	AJ341989 Homo sapi
c	251	18	0.6	804	1	AB039533	AB039533 Pseudomon
	252	18	0.6	837	9	AB049946	AB049946 Homo sapi
	253	18	0.6	885	6	AX327442	AX327442 Sequence
	254	18	0.6	886	9	BC031336	BC031336 Homo sapi
	255	18	0.6	904	9	AF265439	AF265439 Homo sapi
	256	18	0.6	943	9	AF153607	AF153607 Homo sapi
	257	18	0.6	966	1	AF172872	AF172872 Unculture
	258	18	0.6	968	1	AF172915	AF172915 Unculture
	259	18	0.6	972	1	AF172884	AF172884 Unculture
	260	18	0.6	973	1	AF172900	AF172900 Unculture
c	261	18	0.6	1044	6	AX433900	AX433900 Sequence
c	262	18	0.6	1080	10	AF469004	AF469004 Mus muscu
c	263	18	0.6	1105	10	AB073819	AB073819 Mus muscu
	264	18	0.6	1107	4	AF187851	AF187851 Bos tauru
c	265	18	0.6	1179	1	AB014971	AB014971 Ochrobact
c	266	18	0.6	1193	3	AY089425	AY089425 Drosophil
c	267	18	0.6	1281	10	D49684	D49684 Mus musculu
c	268	18	0.6	1287	10	MMBAP37	X78683 M.musculus
	269	18	0.6	1404	10	AY057054	AY057054 Mus muscu
c	270	18	0.6	1530	8	AF009967	AF009967 Candida a
	271	18	0.6	1549	5	AF003532	AF003532 Pleurodel
c	272	18	0.6	1557	9	AF069681	AF069681 Homo sapi
	273	18	0.6	1565	4	AF394922S3	AF394924 Bos tauru
c	274	18	0.6	1579	3	AY058394	AY058394 Drosophil
	275	18	0.6	1595	1	MSGDAPDC	M94109 Mycobacteri
	276	18	0.6	1614	9	AK021626	AK021626 Homo sapi
c	277	18	0.6	1641	9	BC032606	BC032606 Homo sapi
c	278	18	0.6	1645	1	AVU65939	U65939 Azotobacter
c	279	18	0.6	1679	9	AB001835	AB001835 Homo sapi
c	280	18	0.6	1689	6	AR129243	AR129243 Sequence
	281	18	0.6	1691	9	AF055479	AF055479 Homo sapi
	282	18	0.6	1738	4	BTFUT2GEN	X99620 Bos taurus
	283	18	0.6	1807	5	NEWSPA	D63339 Cynops pyrr
	284	18	0.6	1821	10	RATRMCIA	L23128 Rattus norv
c	285	18	0.6	1890	8	DCA249962	AJ249962 Daucus ca
c	286	18	0.6	1963	9	AF051321	AF051321 Homo sapi
	287	18	0.6	2027	6	AX381046	AX381046 Sequence
	288	18	0.6	2080	9	CNS07EG0	AL583864 Human chr
c	289	18	0.6	2104	8	AB049128	AB049128 Cucumis m
	290	18	0.6	2160	5	SST17257	Y17257 Scyliorhinu
c	291	18	0.6	2218	4	CFA388555	AJ388555 Canis fam
	292	18	0.6	2263	9	AK054781	AK054781 Homo sapi
c	293	18	0.6	2283	5	AF033189	AF033189 Gallus ga
c	294	18	0.6	2363	6	E32353	E32353 Plant havin
c	295	18	0.6	2363	8	AF037368	AF037368 Cucumis m
c	296	18	0.6	2457	3	AY118678	AY118678 Drosophil
	297	18	0.6	2641	9	AK091388	AK091388 Homo sapi
c	298	18	0.6	2659	3	DMBRZ2	X54665 D.melanogas
c	299	18	0.6	2682	9	AK001406	AK001406 Homo sapi
	300	18	0.6	2837	9	HSCGG16P5	AJ001218 Homo sapi
c	301	18	0.6	2840	3	DMBRNSZ3	X54664 D.melanogas
c	302	18	0.6	2867	1	SE2XREC	M35138 Plasmid pSE
	303	18	0.6	2979	9	HSL81914	L81914 Homo sapien
c	304	18	0.6	3020	10	MUSAPOIVA	M13966 Mouse apoli
c	305	18	0.6	3033	3	AF257595	AF257595 Drosophil
c	306	18	0.6	3065	3	AF257606	AF257606 Drosophil
c	307	18	0.6	3068	3	AF257590	AF257590 Drosophil
c	308	18	0.6	3068	3	AF257591	AF257591 Drosophil
c	309	18	0.6	3068	3	AF257598	AF257598 Drosophil
c	310	18	0.6	3068	3	AF257600	AF257600 Drosophil
c	311	18	0.6	3068	3	AF257601	AF257601 Drosophil
c	312	18	0.6	3068	3	AF257604	AF257604 Drosophil
c	313	18	0.6	3078	3	AF257592	AF257592 Drosophil
c	314	18	0.6	3078	3	AF257593	AF257593 Drosophil
c	315	18	0.6	3078	3	AF257594	AF257594 Drosophil
c	316	18	0.6	3078	3	AF257596	AF257596 Drosophil
c	317	18	0.6	3078	3	AF257597	AF257597 Drosophil
c	318	18	0.6	3078	3	AF257599	AF257599 Drosophil
c	319	18	0.6	3078	3	AF257602	AF257602 Drosophil
c	320	18	0.6	3078	3	AF257603	AF257603 Drosophil
c	321	18	0.6	3085	3	AF257607	AF257607 Drosophil
c	322	18	0.6	3085	3	AF257608	AF257608 Drosophil
	323	18	0.6	3239	10	BC025581	BC025581 Mus muscu
c	324	18	0.6	3352	8	ATY14404	Y14404 Arabidopsis
c	325	18	0.6	3376	10	BC028975	BC028975 Mus muscu
	326	18	0.6	3486	8	SCYOR378W	Z75286 S.cerevisia
c	327	18	0.6	3647	6	A95922	A95922 Sequence 1
c	328	18	0.6	3647	9	AF151723	AF151723 Homo sapi
c	329	18	0.6	3808	3	DMBRQ1Z1	X54663 D.melanogas
c	330	18	0.6	3869	3	DMU51585	U51585 Drosophila
	331	18	0.6	3916	8	PCH243296	AJ243296 Penicilli
c	332	18	0.6	4003	3	DMBRTNT	X54666 D.melanogas
c	333	18	0.6	4174	3	AY069756	AY069756 Drosophil
	334	18	0.6	4180	6	AX027844	AX027844 Sequence
	335	18	0.6	4180	6	AX027846	AX027846 Sequence
	336	18	0.6	4180	10	MUSCDC25H	L20899 Mouse cell
	337	18	0.6	4338	1	PVU60522	U60522 Paracoccus
	338	18	0.6	4479	4	OCPMA1	X59069 O.cuniculus
c	339	18	0.6	4509	3	DME271350	AJ271350 Drosophil
	340	18	0.6	4948	9	HSM803129	AL831931 Homo sapi
	341	18	0.6	5166	9	AB040880	AB040880 Homo sapi
	342	18	0.6	5370	1	AF112883	AF112883 Pseudomon
	343	18	0.6	5538	9	AB046773	AB046773 Homo sapi
c	344	18	0.6	5710	2	AC013202	AC013202 Drosophil
	345	18	0.6	5788	10	RATMHTLL	M74822 Rat MHC cla
	346	18	0.6	5865	10	RATRMCI	L23127 Rattus norv
c	347	18	0.6	5893	6	AX345761	AX345761 Sequence
	348	18	0.6	5952	6	AX252105	AX252105 Sequence
c	349	18	0.6	7167	6	AX345302	AX345302 Sequence
	350	18	0.6	7432	1	PAU81261	U81261 Pseudomonas
	351	18	0.6	7875	6	E43781	E43781 Genes encod
c	352	18	0.6	9299	6	AX346627	AX346627 Sequence
c	353	18	0.6	10029	1	AE005292	AE005292 Escherich
c	354	18	0.6	10029	1	AE011851	AE011851 Xanthomon
	355	18	0.6	10029	1	AE012501	AE012501 Xanthomon
	356	18	0.6	10201	1	AE005882	AE005882 Caulobact
	357	18	0.6	10368	1	AE002071	AE002071 Deinococc

c	358	18	0.6	10487	1	AE009667	AE009667 Brucella
c	359	18	0.6	10535	1	AE008942	AE008942 Agrobacte
	360	18	0.6	10575	1	AE009238	AE009238 Agrobacte
	361	18	0.6	10608	1	AE009167	AE009167 Agrobacte
	362	18	0.6	10690	1	AE008133	AE008133 Agrobacte
	363	18	0.6	10723	14	AF204178	AF204178 Dengue vi
c	364	18	0.6	10741	1	AE000653	AE000653 Helicobac
c	365	18	0.6	10784	1	AE000200	AE000200 Escherich
	366	18	0.6	11067	1	AE012888	AE012888 Chlorobiu
c	367	18	0.6	11157	1	AE011938	AE011938 Xanthomon
c	368	18	0.6	11176	1	AE007890	AE007890 Agrobacte
	369	18	0.6	11331	1	AE002481	AE002481 Neisseria
c	370	18	0.6	11458	1	AE012204	AE012204 Xanthomon
c	371	18	0.6	11642	1	AE005164	AE005164 Halobacte
	372	18	0.6	11749	1	AE000212	AE000212 Escherich
	373	18	0.6	12250	1	AE012416	AE012416 Xanthomon
	374	18	0.6	12312	1	AE012307	AE012307 Xanthomon
c	375	18	0.6	12398	1	AE001025	AE001025 Archaeogl
c	376	18	0.6	12430	1	AE000370	AE000370 Escherich
c	377	18	0.6	12438	1	AE004044	AE004044 Xylella f
c	378	18	0.6	12986	6	AX251307	AX251307 Sequence
c	379	18	0.6	13086	1	SC1G2	AL136500 Streptomy
	380	18	0.6	13223	1	AE004885	AE004885 Pseudomon
c	381	18	0.6	13949	1	AE005798	AE005798 Caulobact
c	382	18	0.6	13973	8	ATY14403	Y14403 Arabidopsis
c	383	18	0.6	14031	1	AE008376	AE008376 Agrobacte
	384	18	0.6	14106	1	AE007007	AE007007 Mycobacte
	385	18	0.6	14212	2	OSJN00250	AL731607 Oryza sat
	386	18	0.6	14433	1	U67533	U67533 Methanococc
	387	18	0.6	14542	6	AX251844	AX251844 Sequence
	388	18	0.6	14542	6	AX344230	AX344230 Sequence
	389	18	0.6	14542	6	AX348623	AX348623 Sequence
c	390	18	0.6	14789	1	AE004916	AE004916 Pseudomon
	391	18	0.6	16255	1	D90747	D90747 Escherichia
c	392	18	0.6	16259	1	RRU65510	U65510 Rhodospiril
	393	18	0.6	16515	1	AB073929	AB073929 Sporosarc
c	394	18	0.6	16963	8	SPAC16A10	Z97185 S.pombe chr
	395	18	0.6	17212	1	D90746	D90746 Escherichia
c	396	18	0.6	17893	6	AX346267	AX346267 Sequence
c	397	18	0.6	18882	1	D90735	D90735 Escherichia
c	398	18	0.6	19001	3	CEM01A8	Z27081 Caenorhabdi
c	399	18	0.6	19478	3	CET06E8	Z73975 Caenorhabdi
c	400	18	0.6	19650	1	D90736	D90736 Escherichia
c	401	18	0.6	22353	2	AC017269	AC017269 Drosophil
	402	18	0.6	23821	6	E43780	E43780 Genes encod
c	403	18	0.6	25998	9	AF172080	AF172080 Homo sapi
	404	18	0.6	26238	3	U55366	U55366 Caenorhabdi
c	405	18	0.6	29227	3	AF000192	AF000192 Caenorhab
	406	18	0.6	29314	9	AC002455	AC002455 Human cos
c	407	18	0.6	29589	8	SPAC22F8	AL109831 S.pombe c
	408	18	0.6	35516	1	MTCY373	Z73419 Mycobacteri
	409	18	0.6	35594	2	AC019931	AC019931 Drosophil
c	410	18	0.6	36236	9	AL589656	AL589656 Human DNA
	411	18	0.6	36985	1	MSGB1529CS	L78824 Mycobacteri
c	412	18	0.6	37578	3	LMFL7171	AL133435 Leishmani
c	413	18	0.6	38525	3	AF003145	AF003145 Caenorhab
	414	18	0.6	38551	5	TRU414048	AJ414048 Takifugu
c	415	18	0.6	39582	3	DMC17A9	AL009146 Drosophil
c	416	18	0.6	39708	3	CEF35G12	Z46242 Caenorhabdi
	417	18	0.6	41349	2	AC018276	AC018276 Drosophil
	418	18	0.6	42079	3	AC005941	AC005941 Leishmani
c	419	18	0.6	45125	3	AF068714	AF068714 Caenorhab
	420	18	0.6	45597	9	AP000533	AP000533 Homo sapi
c	421	18	0.6	46533	2	AC122120	AC122120 Mus muscu
	422	18	0.6	48622	8	AB008266	AB008266 Arabidops
c	423	18	0.6	49600	8	AP004342	AP004342 Oryza sat
c	424	18	0.6	50960	2	AC017313	AC017313 Drosophil
c	425	18	0.6	51111	2	AC100579	AC100579 Mus muscu
	426	18	0.6	51661	8	AP003864	AP003864 Oryza sat
	427	18	0.6	52586	9	AF258547	AF258547 Homo sapi
	428	18	0.6	53931	2	AC100941	AC100941 Mus muscu
	429	18	0.6	54299	2	AC124645	AC124645 Mus muscu
c	430	18	0.6	55175	1	ECU28375	U28375 Escherichia

c	431	18	0.6	55739	8	AP000389	AP000389 Arabidops
c	432	18	0.6	55892	5	AF013614	AF013614 Fugu rubr
c	433	18	0.6	56374	2	AC099855	AC099855 Mus muscu
	434	18	0.6	60009	2	AC023436	AC023436 Homo sapi
c	435	18	0.6	60022	2	AC102916	AC102916 Mus muscu
	436	18	0.6	60196	2	AC124052	AC124052 Mus muscu
c	437	18	0.6	60286	2	AC100131	AC100131 Mus muscu
	438	18	0.6	61575	2	AC100040	AC100040 Mus muscu
c	439	18	0.6	61575	2	AC100040	AC100040 Mus muscu
c	440	18	0.6	62838	2	AC061995	AC061995 Homo sapi
	441	18	0.6	63578	9	AC109356	AC109356 Homo sapi
	442	18	0.6	64089	2	AC100165	AC100165 Mus muscu
	443	18	0.6	64295	2	AC100066	AC100066 Mus muscu
c	444	18	0.6	64391	2	AC090856	AC090856 Homo sapi
	445	18	0.6	64829	2	AC100338	AC100338 Mus muscu
c	446	18	0.6	65547	8	AY013245	AY013245 Oryza sat
	447	18	0.6	66281	9	AL138710	AL138710 Human DNA
c	448	18	0.6	67555	2	AC100398	AC100398 Mus muscu
	449	18	0.6	67859	2	AC101982	AC101982 Mus muscu
c	450	18	0.6	68904	2	AC027370	AC027370 Homo sapi
c	451	18	0.6	71821	2	AC091584	AC091584 Homo sapi
c	452	18	0.6	72924	2	AC025930	AC025930 Homo sapi
c	453	18	0.6	76179	2	AC012976	AC012976 Drosophil
c	454	18	0.6	76529	8	H0423H10	AL442112 Oryza sat
	455	18	0.6	76606	9	AL162588	AL162588 Human DNA
	456	18	0.6	78276	3	AC024761	AC024761 Caenorhab
	457	18	0.6	80308	2	AC105512	AC105512 Rattus no
c	458	18	0.6	80622	2	AC109072	AC109072 Rattus no
c	459	18	0.6	82912	3	AC004574	AC004574 Drosophil
c	460	18	0.6	83013	2	AC111418	AC111418 Rattus no
	461	18	0.6	84243	2	AC122115	AC122115 Homo sapi
	462	18	0.6	84514	10	AL627264	AL627264 Mouse DNA
c	463	18	0.6	84702	8	AB018114	AB018114 Arabidops
	464	18	0.6	85995	2	AC027404	AC027404 Homo sapi
c	465	18	0.6	86530	2	AC111753	AC111753 Rattus no
	466	18	0.6	88318	8	AP002842	AP002842 Oryza sat
c	467	18	0.6	88937	2	AC097874	AC097874 Rattus no
	468	18	0.6	89818	9	AC002126	AC002126 Homo sapi
	469	18	0.6	91040	2	AC017292	AC017292 Drosophil
	470	18	0.6	91573	9	AC012372	AC012372 Homo sapi
c	471	18	0.6	93234	8	AC009465	AC009465 Arabidops
	472	18	0.6	93398	2	AP001863	AP001863 Homo sapi
c	473	18	0.6	95009	8	ATF25G13	AL079349 Arabidops
	474	18	0.6	95129	9	AC003099	AC003099 Homo sapi
	475	18	0.6	95597	2	AC014022	AC014022 Drosophil
	476	18	0.6	95746	2	AC020248	AC020248 Drosophil
	477	18	0.6	96006	2	AL139224	AL139224 Homo sapi
	478	18	0.6	96006	9	HSJ590P13	AL109743 Human DNA
c	479	18	0.6	96485	2	AC130635	AC130635 Rattus no
c	480	18	0.6	96705	2	AC098198	AC098198 Rattus no
c	481	18	0.6	96745	2	AC074219	AC074219 Mus muscu
	482	18	0.6	97033	2	AC015146	AC015146 Drosophil
	483	18	0.6	97943	9	HSAC002467	AC002467 Human BAC
c	484	18	0.6	98062	2	RN516O10	AL603719 Rattus no
	485	18	0.6	99504	2	AC098401	AC098401 Rattus no
c	486	18	0.6	99802	9	AL158169	AL158169 Human DNA
c	487	18	0.6	99940	9	AC079246	AC079246 Homo sapi
	488	18	0.6	100786	2	AC013972	AC013972 Drosophil
	489	18	0.6	102378	2	AP003736	AP003736 Oryza sat
c	490	18	0.6	104577	5	AL713869	AL713869 Zebrafish
c	491	18	0.6	104726	9	AC008087	AC008087 Homo sapi
	492	18	0.6	105362	9	AC012370	AC012370 Homo sapi
c	493	18	0.6	107037	9	HSJ678E16	AL122004 Human DNA
	494	18	0.6	108232	2	AC099662	AC099662 Rattus no
c	495	18	0.6	108388	9	AP001821	AP001821 Homo sapi
c	496	18	0.6	108964	8	AC073391	AC073391 Oryza sat
	497	18	0.6	110000	2	LMFLCHR32_19	Continuation (20 o
	498	18	0.6	110000	2	LMFLCHR32_20	Continuation (21 o
c	499	18	0.6	110000	2	LMFLCHR34_14	Continuation (15 o
	500	18	0.6	110000	2	TBBCHR1A_02	Continuation (3 of
	501	18	0.6	110000	3	AC125735_0	AC125735 leishmani
	502	18	0.6	110000	3	AC125735_1	Continuation (2 of
c	503	18	0.6	110338	3	AC006066	AC006066 Drosophil

	504	18	0.6	110611	8	AC022520	AC022520 Arabidops
c	505	18	0.6	110721	2	AC098764	AC098764 Rattus no
c	506	18	0.6	111172	2	AP005293	AP005293 Oryza sat
c	507	18	0.6	111469	2	AC022470	AC022470 Homo sapi
c	508	18	0.6	111634	9	AC025755	AC025755 Homo sapi
c	509	18	0.6	111837	9	AC069383	AC069383 Homo sapi
	510	18	0.6	113031	9	AL445672	AL445672 Human DNA
c	511	18	0.6	113639	9	HSJ744A17	AL096862 Human DNA
	512	18	0.6	113646	2	AC106265	AC106265 Rattus no
c	513	18	0.6	113988	9	AL133546	AL133546 Human DNA
c	514	18	0.6	114070	2	AP005243	AP005243 Oryza sat
c	515	18	0.6	114584	2	AF252827	AF252827 Homo sapi
	516	18	0.6	116004	9	AF258545	AF258545 Homo sapi
	517	18	0.6	116801	2	AL390054	AL390054 Homo sapi
c	518	18	0.6	117716	2	AC094474	AC094474 Rattus no
	519	18	0.6	117779	10	AL732293	AL732293 Mouse DNA
c	520	18	0.6	119061	2	AL669909	AL669909 Mus muscu
c	521	18	0.6	119066	2	AC026377	AC026377 Mus muscu
c	522	18	0.6	119243	2	AC112576	AC112576 Rattus no
	523	18	0.6	119295	2	AC127427	AC127427 Magnaport
	524	18	0.6	119797	2	AC013236	AC013236 Drosophil
c	525	18	0.6	120030	9	AC002075	AC002075 Homo sapi
c	526	18	0.6	120159	2	AC120509	AC120509 Oryza sat
c	527	18	0.6	120641	2	AC094693	AC094693 Rattus no
	528	18	0.6	121019	2	AC102708	AC102708 Mus muscu
	529	18	0.6	122648	2	AC026025	AC026025 Homo sapi
	530	18	0.6	123427	8	AC091732	AC091732 Oryza sat
c	531	18	0.6	124093	2	AP005575	AP005575 Oryza sat
	532	18	0.6	124989	2	CNS08C8P	AL731885 Oryza sat
c	533	18	0.6	126155	2	AC125631	AC125631 Rattus no
c	534	18	0.6	126175	2	AC127717	AC127717 Rattus no
	535	18	0.6	126315	2	AP004116	AP004116 Oryza sat
	536	18	0.6	126659	8	OSJN00015	AL606455 Oryza sat
c	537	18	0.6	127204	2	AC122114	AC122114 Atelerix
	538	18	0.6	127423	2	AC117018	AC117018 Rattus no
c	539	18	0.6	127432	2	AC129056	AC129056 Rattus no
	540	18	0.6	128210	2	AP005643	AP005643 Oryza sat
c	541	18	0.6	128342	9	AL627313	AL627313 Human DNA
	542	18	0.6	129104	2	CNS08C92	AL732379 Oryza sat
c	543	18	0.6	130235	8	AC008007	AC008007 Genomic s
c	544	18	0.6	130266	9	AL596220	AL596220 Human DNA
	545	18	0.6	130941	2	AC105770	AC105770 Oryza sat
	546	18	0.6	132104	2	AC073672	AC073672 Mus muscu
	547	18	0.6	132384	9	AC116347	AC116347 Homo sapi
	548	18	0.6	132755	9	AC008670	AC008670 Homo sapi
c	549	18	0.6	132927	8	AC079890	AC079890 Oryza sat
c	550	18	0.6	133028	8	AP003447	AP003447 Oryza sat
c	551	18	0.6	133461	2	AC025968	AC025968 Homo sapi
	552	18	0.6	134617	2	AC115001	AC115001 Mus muscu
	553	18	0.6	134725	2	AC114389	AC114389 Rattus no
	554	18	0.6	135357	2	AP004769	AP004769 Oryza sat
	555	18	0.6	135743	2	AC123302	AC123302 Rattus no
c	556	18	0.6	136356	9	AC040166	AC040166 Homo sapi
	557	18	0.6	136906	2	RN86I20	AL603805 Rattus no
	558	18	0.6	137071	2	AC119540	AC119540 Rattus no
	559	18	0.6	138633	9	AP003078	AP003078 Homo sapi
c	560	18	0.6	138893	2	AC107593	AC107593 Rattus no
	561	18	0.6	138904	2	AP003928	AP003928 Oryza sat
c	562	18	0.6	138939	9	AC119675	AC119675 Homo sapi
c	563	18	0.6	139005	2	AF286112	AF286112 Homo sapi
c	564	18	0.6	139468	8	AC090441	AC090441 Oryza sat
c	565	18	0.6	139581	2	AP005055	AP005055 Oryza sat
	566	18	0.6	139814	2	AP003964	AP003964 Oryza sat
	567	18	0.6	140191	2	AC109434	AC109434 Rattus no
c	568	18	0.6	140474	2	AC116947	AC116947 Tetraodon
c	569	18	0.6	140583	2	AP005413	AP005413 Oryza sat
c	570	18	0.6	140714	2	RN374E16	AL603726 Rattus no
	571	18	0.6	140774	2	AC108904	AC108904 Felis cat
c	572	18	0.6	141386	2	AC119454	AC119454 Rattus no
c	573	18	0.6	141777	9	AC011738	AC011738 Homo sapi
	574	18	0.6	142164	2	AC115193	AC115193 Rattus no
c	575	18	0.6	142907	2	AC091330	AC091330 Trypanoso
c	576	18	0.6	143252	2	CNS08CB5	AL845343 Oryza sat
	577	18	0.6	143969	8	AP002839	AP002839 Oryza sat
	578	18	0.6	144191	8	AC079874	AC079874 Oryza sat
c	579	18	0.6	144321	2	AC130603	AC130603 Oryza sat
	580	18	0.6	144538	9	CNS07ECX	AL445074 Human chr
	581	18	0.6	144784	9	AC009224	AC009224 Homo sapi
	582	18	0.6	144983	2	AC105284	AC105284 Homo sapi
c	583	18	0.6	145242	9	AC013457	AC013457 Homo sapi
c	584	18	0.6	145299	10	AL603826	AL603826 Mouse DNA
	585	18	0.6	148370	10	AC091514	AC091514 Rattus no
	586	18	0.6	148375	2	AC118789	AC118789 Rattus no
c	587	18	0.6	148381	8	AP003253	AP003253 Oryza sat
	588	18	0.6	148507	2	AC011263	AC011263 Homo sapi
	589	18	0.6	148704	8	AC124213	AC124213 Genomic s
c	590	18	0.6	148750	9	AL596225	AL596225 Human DNA
c	591	18	0.6	149347	8	AC073392	AC073392 Oryza sat
c	592	18	0.6	149945	2	AC124146	AC124146 Rattus no
c	593	18	0.6	150113	9	AC015988	AC015988 Homo sapi
	594	18	0.6	150381	8	AP003794	AP003794 Oryza sat
c	595	18	0.6	150690	9	AC090402	AC090402 Homo sapi
	596	18	0.6	151243	2	AC094894	AC094894 Rattus no
	597	18	0.6	151551	2	AC021291	AC021291 Homo sapi
c	598	18	0.6	151847	2	AC022792	AC022792 Homo sapi
	599	18	0.6	152668	2	AC087623	AC087623 Homo sapi
	600	18	0.6	153297	2	AC099458	AC099458 Rattus no
c	601	18	0.6	153297	2	AC099458	AC099458 Rattus no
	602	18	0.6	153448	2	AC112088	AC112088 Rattus no
	603	18	0.6	153666	8	AP003287	AP003287 Oryza sat
	604	18	0.6	153789	2	AP004812	AP004812 Oryza sat
c	605	18	0.6	153954	9	AL354862	AL354862 Human DNA
c	606	18	0.6	154242	2	AC073373	AC073373 Mus muscu
	607	18	0.6	154614	10	AC121586	AC121586 Mus muscu
c	608	18	0.6	154832	2	AC119142	AC119142 Rattus no
	609	18	0.6	155060	2	AP005591	AP005591 Oryza sat
	610	18	0.6	155152	2	AP004858	AP004858 Oryza sat
c	611	18	0.6	155276	2	AC018410	AC018410 Homo sapi
c	612	18	0.6	156127	2	AP005156	AP005156 Oryza sat
	613	18	0.6	156143	2	AC128712	AC128712 Homo sapi
c	614	18	0.6	156440	2	AC111368	AC111368 Rattus no
c	615	18	0.6	156472	2	AL732597	AL732597 Mus muscu
c	616	18	0.6	156763	9	AC005901	AC005901 Homo sapi
c	617	18	0.6	157021	2	AC129066	AC129066 Didelphis
	618	18	0.6	157451	2	AC120275	AC120275 Rattus no
c	619	18	0.6	157926	2	AC096047	AC096047 Rattus no
c	620	18	0.6	158105	2	AC127408	AC127408 Rattus no
c	621	18	0.6	158633	2	AC015472	AC015472 Homo sapi
c	622	18	0.6	158641	2	AC102786	AC102786 Mus muscu
c	623	18	0.6	159191	2	AP005099	AP005099 Oryza sat
c	624	18	0.6	159538	2	AC129676	AC129676 Homo sapi
c	625	18	0.6	159611	9	AC026475	AC026475 Homo sapi
	626	18	0.6	159700	2	AC010822	AC010822 Homo sapi
	627	18	0.6	159712	9	AP004219	AP004219 Homo sapi
c	628	18	0.6	159784	2	AC061996	AC061996 Homo sapi
	629	18	0.6	159942	9	AC025018	AC025018 Homo sapi
	630	18	0.6	160026	2	AC021437	AC021437 Homo sapi
	631	18	0.6	160030	2	CNS07YPB	AL713942 Oryza sat
c	632	18	0.6	160551	3	AC008320	AC008320 Drosophil
	633	18	0.6	160573	9	AC119676	AC119676 Homo sapi
c	634	18	0.6	160635	2	AC129997	AC129997 Rattus no
c	635	18	0.6	160796	9	AL499604	AL499604 Human DNA
	636	18	0.6	160831	2	AC025015	AC025015 Homo sapi
	637	18	0.6	161313	2	AC107846	AC107846 Mus muscu
	638	18	0.6	161428	2	AC126925	AC126925 Canis fam
c	639	18	0.6	161873	8	AC090056	AC090056 Oryza sat
c	640	18	0.6	161976	2	AC012415	AC012415 Homo sapi
c	641	18	0.6	162060	2	AC007515	AC007515 Homo sapi
c	642	18	0.6	162560	2	AL358874	AL358874 Homo sapi
c	643	18	0.6	162845	2	AC101829	AC101829 Mus muscu
c	644	18	0.6	163158	2	AC111278	AC111278 Rattus no
	645	18	0.6	164018	9	AL354868	AL354868 Human DNA
	646	18	0.6	164810	2	AC115069	AC115069 Mus muscu
	647	18	0.6	164911	2	AC119581	AC119581 Rattus no
c	648	18	0.6	165341	2	AC092037	AC092037 Homo sapi
	649	18	0.6	165459	2	AC121653	AC121653 Rattus no

	650	18	0.6	165901	9	AP003474	AP003474 Homo sapi
c	651	18	0.6	166013	2	AC016328	AC016328 Homo sapi
	652	18	0.6	166133	2	AC106907	AC106907 Rattus no
	653	18	0.6	166352	3	AC010069	AC010069 Drosophil
c	654	18	0.6	166543	3	AC005813	AC005813 Drosophil
	655	18	0.6	166569	2	AC106496	AC106496 Rattus no
c	656	18	0.6	166592	2	AC126909	AC126909 Rattus no
c	657	18	0.6	166899	8	AP003350	AP003350 Oryza sat
c	658	18	0.6	167102	9	AC084017	AC084017 Homo sapi
c	659	18	0.6	167289	2	AC093187	AC093187 Papio cyn
c	660	18	0.6	167617	2	AC091403	AC091403 Sus scrof
c	661	18	0.6	167630	2	AC105064	AC105064 Mus muscu
c	662	18	0.6	167694	9	AC113137	AC113137 Homo sapi
	663	18	0.6	167888	2	AC129805	AC129805 Homo sapi
	664	18	0.6	167914	2	AC019268	AC019268 Homo sapi
	665	18	0.6	168187	2	AC094137	AC094137 Rattus no
c	666	18	0.6	168505	2	AC107142	AC107142 Rattus no
	667	18	0.6	168562	2	AC107012	AC107012 Rattus no
c	668	18	0.6	169249	2	AC022815	AC022815 Homo sapi
c	669	18	0.6	170025	2	OSJN00031	AL606590 Oryza sat
	670	18	0.6	170082	2	AC129774	AC129774 Mus muscu
c	671	18	0.6	170341	3	AC013423	AC013423 Drosophil
	672	18	0.6	170369	2	AF214633	AF214633 Homo sapi
	673	18	0.6	170523	2	AC053526	AC053526 Homo sapi
c	674	18	0.6	170931	2	AC096400	AC096400 Rattus no
	675	18	0.6	170959	2	AC022197	AC022197 Homo sapi
	676	18	0.6	171076	2	AC025560	AC025560 Homo sapi
	677	18	0.6	171114	9	AC092757	AC092757 Homo sapi
	678	18	0.6	171168	2	AP005385	AP005385 Oryza sat
c	679	18	0.6	171179	2	AL162722	AL162722 Homo sapi
	680	18	0.6	171427	9	AC021590	AC021590 Homo sapi
	681	18	0.6	171710	2	AC068049	AC068049 Homo sapi
	682	18	0.6	172024	10	AL606904	AL606904 Mouse DNA
c	683	18	0.6	172078	2	AC019359	AC019359 Homo sapi
c	684	18	0.6	172154	2	AC018396	AC018396 Homo sapi
	685	18	0.6	172277	2	AC099233	AC099233 Rattus no
c	686	18	0.6	172279	2	AC107529	AC107529 Rattus no
c	687	18	0.6	172349	9	AC019185	AC019185 Homo sapi
c	688	18	0.6	172367	2	AC118989	AC118989 Canis fam
	689	18	0.6	172512	9	AC016903	AC016903 Homo sapi
	690	18	0.6	173645	2	AC123526	AC123526 Oryza sat
	691	18	0.6	173691	2	AC087462	AC087462 Homo sapi
c	692	18	0.6	173769	9	AC113933	AC113933 Homo sapi
	693	18	0.6	173797	3	AC105898	AC105898 Drosophil
c	694	18	0.6	173932	2	AC102837	AC102837 Mus muscu
	695	18	0.6	174247	2	AC095090	AC095090 Rattus no
c	696	18	0.6	174380	2	AP003720	AP003720 Homo sapi
	697	18	0.6	174419	2	AC119060	AC119060 Bos tauru
c	698	18	0.6	174480	2	AC026069	AC026069 Homo sapi
c	699	18	0.6	174602	2	AC106552	AC106552 Rattus no
c	700	18	0.6	174616	3	AC010016	AC010016 Drosophil
	701	18	0.6	174793	9	CNS05TE5	AL358334 Human chr
c	702	18	0.6	174877	2	AP004818	AP004818 Oryza sat
c	703	18	0.6	174919	2	AC120741	AC120741 Rattus no
	704	18	0.6	175115	3	AC105899	AC105899 Drosophil
c	705	18	0.6	175269	2	AC120762	AC120762 Rattus no
c	706	18	0.6	175430	2	AC082646	AC082646 Homo sapi
	707	18	0.6	175456	2	AC118909	AC118909 Rattus no
	708	18	0.6	175911	2	AC131355	AC131355 Rattus no
c	709	18	0.6	175928	2	AC129750	AC129750 Rattus no
	710	18	0.6	175947	8	AP003435	AP003435 Oryza sat
	711	18	0.6	176010	2	AC069125	AC069125 Homo sapi
c	712	18	0.6	176095	2	AC019346	AC019346 Homo sapi
	713	18	0.6	176194	9	AL691486	AL691486 Human DNA
c	714	18	0.6	176580	2	AC102248	AC102248 Mus muscu
c	715	18	0.6	176580	2	AP003634	AP003634 Oryza sat
	716	18	0.6	176642	10	AC121987	AC121987 Mus muscu
	717	18	0.6	176915	2	AL844176	AL844176 Mus muscu
	718	18	0.6	176928	9	AC084239	AC084239 Homo sapi
c	719	18	0.6	176971	2	AC027455	AC027455 Homo sapi
	720	18	0.6	176995	9	AC022735	AC022735 Homo sapi
	721	18	0.6	177018	2	AC106357	AC106357 Rattus no
c	722	18	0.6	177096	3	AC107326	AC107326 Drosophil
	723	18	0.6	177320	9	AC104441	AC104441 Homo sapi
c	724	18	0.6	177688	9	CNS01DT5	AL132640 Human chr
	725	18	0.6	177782	9	AC117532	AC117532 Homo sapi
c	726	18	0.6	177865	9	AC098483	AC098483 Homo sapi
	727	18	0.6	178199	2	AC018772	AC018772 Homo sapi
c	728	18	0.6	178714	2	AC017008	AC017008 Homo sapi
c	729	18	0.6	178959	2	AP004809	AP004809 Oryza sat
c	730	18	0.6	179343	2	AC121904	AC121904 Mus muscu
	731	18	0.6	179683	2	AC101721	AC101721 Mus muscu
c	732	18	0.6	179683	2	AC101721	AC101721 Mus muscu
c	733	18	0.6	179796	2	AC098948	AC098948 Rattus no
	734	18	0.6	179950	2	AC111288	AC111288 Rattus no
c	735	18	0.6	180009	2	AC023259	AC023259 Homo sapi
c	736	18	0.6	180017	2	AC068669	AC068669 Homo sapi
c	737	18	0.6	180773	9	AC092969	AC092969 Homo sapi
	738	18	0.6	180890	2	AC018402	AC018402 Homo sapi
c	739	18	0.6	180915	2	AC119042	AC119042 Homo sapi
c	740	18	0.6	181009	3	AC007575	AC007575 Drosophil
c	741	18	0.6	181486	2	AC116503	AC116503 Mus muscu
	742	18	0.6	181607	9	AC021581	AC021581 Homo sapi
	743	18	0.6	181780	2	AC116814	AC116814 Mus muscu
	744	18	0.6	181784	2	AC117172	AC117172 Rattus no
	745	18	0.6	181927	3	AC010573	AC010573 Drosophil
c	746	18	0.6	182322	9	AC092379	AC092379 Homo sapi
c	747	18	0.6	182339	2	AC092761	AC092761 Papio cyn
	748	18	0.6	182528	9	AC020743	AC020743 Homo sapi
	749	18	0.6	182540	2	AC016916	AC016916 Homo sapi
c	750	18	0.6	182622	2	AC117173	AC117173 Rattus no
c	751	18	0.6	183043	2	AC106096	AC106096 Rattus no
c	752	18	0.6	183067	9	AC040914	AC040914 Homo sapi
	753	18	0.6	183193	2	AL844554	AL844554 Mus muscu
c	754	18	0.6	183249	9	AC004828	AC004828 Homo sapi
c	755	18	0.6	183338	2	AC069514	AC069514 Homo sapi
c	756	18	0.6	183370	3	AC008213	AC008213 Drosophil
	757	18	0.6	183719	2	AC068842	AC068842 Homo sapi
	758	18	0.6	184172	2	AC130589	AC130589 Rattus no
	759	18	0.6	184267	2	AC078796	AC078796 Homo sapi
	760	18	0.6	184315	2	AP005621	AP005621 Oryza sat
	761	18	0.6	184509	2	AC113237	AC113237 Canis fam
c	762	18	0.6	184582	9	AC010207	AC010207 Homo sapi
c	763	18	0.6	184888	2	AP004795	AP004795 Oryza sat
c	764	18	0.6	185096	9	AL355140	AL355140 Human DNA
	765	18	0.6	185165	9	AC008732	AC008732 Homo sapi
c	766	18	0.6	185311	9	AL355586	AL355586 Human DNA
	767	18	0.6	185563	2	AC112284	AC112284 Rattus no
c	768	18	0.6	185624	9	AC108162	AC108162 Homo sapi
	769	18	0.6	185937	9	AC107896	AC107896 Homo sapi
	770	18	0.6	185948	2	AC098901	AC098901 Rattus no
	771	18	0.6	185952	9	AC005906	AC005906 Homo sapi
c	772	18	0.6	186359	2	AC111041	AC111041 Mus muscu
	773	18	0.6	186416	2	AC011226	AC011226 Homo sapi
c	774	18	0.6	186714	3	AC023688	AC023688 Drosophil
	775	18	0.6	186973	2	AC124932	AC124932 Rattus no
c	776	18	0.6	187193	9	AC007222	AC007222 Homo sapi
c	777	18	0.6	187600	2	AC092515	AC092515 Papio cyn
	778	18	0.6	187734	9	AC025918	AC025918 Homo sapi
c	779	18	0.6	187908	9	HS495O10	AL031121 Human DNA
	780	18	0.6	188072	2	AC009040	AC009040 Homo sapi
c	781	18	0.6	188494	2	AC126154	AC126154 Rattus no
	782	18	0.6	188509	2	AC115933	AC115933 Mus muscu
c	783	18	0.6	188637	2	AL845306	AL845306 Danio rer
c	784	18	0.6	188716	2	AC015886	AC015886 Mus muscu
c	785	18	0.6	188740	2	AC024045	AC024045 Homo sapi
	786	18	0.6	189008	2	AP001491	AP001491 Homo sapi
	787	18	0.6	189179	2	AC090265	AC090265 Homo sapi
	788	18	0.6	189466	9	AC018887	AC018887 Homo sapi
	789	18	0.6	189722	9	AC067793	AC067793 Homo sapi
	790	18	0.6	189771	2	AC112921	AC112921 Homo sapi
	791	18	0.6	190017	9	AC090589	AC090589 Homo sapi
	792	18	0.6	190091	2	AC123953	AC123953 Mus muscu
	793	18	0.6	191028	8	PPU38804	U38804 Porphyra pu
c	794	18	0.6	191119	9	AC009957	AC009957 Homo sapi
	795	18	0.6	191556	2	AC026117	AC026117 Homo sapi

	796	18	0.6	191625	2	AC118038 AC118038 Mus muscu
c	797	18	0.6	191957	9	AC015845 AC015845 Homo sapi
	798	18	0.6	192321	2	AC119324 AC119324 Rattus no
c	799	18	0.6	192387	9	CNS01DV0 AL133444 Human chr
c	800	18	0.6	192966	9	AL353813 AL353813 Human DNA
	801	18	0.6	192968	2	AC099655 AC099655 Rattus no
c	802	18	0.6	193046	2	AC062034 AC062034 Homo sapi
c	803	18	0.6	193260	2	AC025598 AC025598 Homo sapi
c	804	18	0.6	194329	2	AC116404 AC116404 Mus muscu
	805	18	0.6	194646	2	AC090129 AC090129 Homo sapi
c	806	18	0.6	194790	3	AC009840 AC009840 Drosophil
	807	18	0.6	194905	2	AL603764 AL603764 Homo sapi
	808	18	0.6	195270	9	AC096534 AC096534 Homo sapi
	809	18	0.6	195296	2	AC026606 AC026606 Homo sapi
	810	18	0.6	195348	2	AC127081 AC127081 Rattus no
c	811	18	0.6	196259	2	AL135915 AL135915 Homo sapi
c	812	18	0.6	196544	9	AC015974 AC015974 Homo sapi
	813	18	0.6	196726	2	AC126723 AC126723 Rattus no
c	814	18	0.6	197078	2	AC117382 AC117382 Homo sapi
c	815	18	0.6	197127	2	AC098253 AC098253 Rattus no
c	816	18	0.6	197437	2	AL772303 AL772303 Mus muscu
	817	18	0.6	197597	3	AC011253 AC011253 Drosophil
c	818	18	0.6	197751	2	AC073290 AC073290 Mus muscu
c	819	18	0.6	198226	8	ATCHRIV43 AL161543 Arabidops
c	820	18	0.6	198348	2	AL772277 AL772277 Mus muscu
c	821	18	0.6	198453	9	AC093889 AC093889 Homo sapi
c	822	18	0.6	199133	2	AC122212 AC122212 Mus muscu
c	823	18	0.6	199280	8	ATCHRIV35 AL161535 Arabidops
c	824	18	0.6	200441	2	AC117150 AC117150 Rattus no
	825	18	0.6	200966	10	AL596108 AL596108 Mouse DNA
	826	18	0.6	201652	2	AC117133 AC117133 Rattus no
	827	18	0.6	201766	2	AL845260 AL845260 Mus muscu
c	828	18	0.6	201989	2	AC073169 AC073169 Homo sapi
	829	18	0.6	202685	2	AC127053 AC127053 Rattus no
	830	18	0.6	203286	10	AL606829 AL606829 Mouse DNA
	831	18	0.6	203914	2	AC130005 AC130005 Homo sapi
c	832	18	0.6	204317	2	AC117614 AC117614 Mus muscu
c	833	18	0.6	204550	2	AL844199 AL844199 Mus muscu
c	834	18	0.6	204708	2	AC091617 AC091617 Rattus no
	835	18	0.6	204775	2	AC120549 AC120549 Mus muscu
	836	18	0.6	205397	2	AC129193 AC129193 Mus muscu
	837	18	0.6	205453	2	AC079556 AC079556 Mus muscu
	838	18	0.6	205577	2	AC011237 AC011237 Homo sapi
	839	18	0.6	205749	10	AL671187 AL671187 Mouse DNA
	840	18	0.6	206082	2	AC126936 AC126936 Mus muscu
	841	18	0.6	206125	10	AL672219 AL672219 Mouse DNA
	842	18	0.6	208197	9	AC009107 AC009107 Homo sapi
	843	18	0.6	208327	2	AC119061 AC119061 Bos tauru
	844	18	0.6	208330	2	AC022327 AC022327 Mus muscu
c	845	18	0.6	208804	2	AC117025 AC117025 Rattus no
c	846	18	0.6	209164	8	ATFCA5 Z97340 Arabidopsis
	847	18	0.6	209216	2	AC117126 AC117126 Rattus no
	848	18	0.6	209600	2	AC022452 AC022452 Mus muscu
c	849	18	0.6	210336	2	AC022171 AC022171 Homo sapi
	850	18	0.6	210376	2	AL136118 AL136118 Homo sapi
c	851	18	0.6	210514	10	AC117199 AC117199 Mus muscu
c	852	18	0.6	211084	2	AC092084 AC092084 Bos tauru
	853	18	0.6	211217	2	AC123840 AC123840 Mus muscu
c	854	18	0.6	212085	10	AL663030 AL663030 Mouse DNA
	855	18	0.6	213251	2	AC113290 AC113290 Mus muscu
	856	18	0.6	213452	2	AL663090 AL663090 Mus muscu
	857	18	0.6	214166	9	AC009660 AC009660 Homo sapi
c	858	18	0.6	214967	2	AC130539 AC130539 Mus muscu
c	859	18	0.6	215050	1	AL646057 AL646057 Ralstonia
c	860	18	0.6	216776	10	AL669920 AL669920 Mouse DNA
	861	18	0.6	217017	2	AC124385 AC124385 Mus muscu
	862	18	0.6	217130	2	AC123838 AC123838 Mus muscu
	863	18	0.6	217352	10	AL645784 AL645784 Mouse DNA
	864	18	0.6	217420	1	BSUB0017 Z99120 Bacillus su
c	865	18	0.6	218193	2	AC121608 AC121608 Mus muscu
	866	18	0.6	218209	2	AC099425 AC099425 Rattus no
	867	18	0.6	218298	2	AC116389 AC116389 Mus muscu
c	868	18	0.6	219709	2	AC091360 AC091360 Rattus no
	869	18	0.6	220633	9	HUU91321 U91321 Human Chrom
c	870	18	0.6	221883	2	AC104752 AC104752 Mus muscu
	871	18	0.6	221912	10	AL672003 AL672003 Mouse DNA
c	872	18	0.6	222618	2	AC103335 AC103335 Rattus no
	873	18	0.6	223280	2	AC130456 AC130456 Homo sapi
c	874	18	0.6	223326	2	AC127567 AC127567 Mus muscu
	875	18	0.6	223898	2	AC124604 AC124604 Mus muscu
	876	18	0.6	224068	9	CNS01DUJ AL133246 BAC seque
c	877	18	0.6	224757	2	AC122913 AC122913 Mus muscu
c	878	18	0.6	224799	2	AC122472 AC122472 Mus muscu
	879	18	0.6	225083	10	AL596382 AL596382 Mouse DNA
	880	18	0.6	225136	2	AC122866 AC122866 Mus muscu
	881	18	0.6	227538	10	AC002397 AC002397 Mouse chr
c	882	18	0.6	228263	2	AC079930 AC079930 Homo sapi
	883	18	0.6	228516	9	AC002523 AC002523 Homo sapi
	884	18	0.6	230458	2	AL731828 AL731828 Mus muscu
	885	18	0.6	232904	2	AL732546 AL732546 Mus muscu
c	886	18	0.6	232990	2	AC102545 AC102545 Mus muscu
c	887	18	0.6	234189	2	AC125040 AC125040 Mus muscu
	888	18	0.6	238016	2	AL773538 AL773538 Mus muscu
c	889	18	0.6	240871	2	AC122568 AC122568 Rattus no
	890	18	0.6	241886	10	AC027285 AC027285 Mus muscu
	891	18	0.6	242030	2	AC122673 AC122673 Rattus no
c	892	18	0.6	244164	3	AE003548 AE003548 Drosophil
c	893	18	0.6	247850	10	AJ421478 AJ421478 Mus muscu
	894	18	0.6	250991	2	AC127270 AC127270 Mus muscu
c	895	18	0.6	256242	3	AE003756 AE003756 Drosophil
	896	18	0.6	256373	2	AC020870 AC020870 Mus muscu
c	897	18	0.6	260050	1	RME603643 AL603643 Rhizobium
c	898	18	0.6	261156	2	AC105829 AC105829 Rattus no
c	899	18	0.6	263744	2	AF401201 AF401201 Mus muscu
c	900	18	0.6	263859	2	AC025353 AC025353 Mus muscu
c	901	18	0.6	264853	3	AE003644 AE003644 Drosophil
	902	18	0.6	269081	2	AC068493 AC068493 Mus muscu
c	903	18	0.6	271519	2	AC020887 AC020887 Mus muscu
c	904	18	0.6	273785	1	SME591793 AL591793 Sinorhizo
c	905	18	0.6	276829	2	AC006741 AC006741 Caenorhab
	906	18	0.6	278708	3	AE003535 AE003535 Drosophil
c	907	18	0.6	291622	2	AC006796 AC006796 Caenorhab
c	908	18	0.6	291829	3	AE003525 AE003525 Drosophil
c	909	18	0.6	293431	2	PFMAL13P4 AL049181 Plasmodiu
c	910	18	0.6	294914	3	AE003603 AE003603 Drosophil
c	911	18	0.6	296497	3	AE003421 AE003421 Drosophil
c	912	18	0.6	299782	2	AC006844 AC006844 Caenorhab
c	913	18	0.6	300962	3	DROSADH04 AE003410 Drosophil
	914	18	0.6	302797	3	AE003488 AE003488 Drosophil
c	915	18	0.6	306321	3	AE003585 AE003585 Drosophil
	916	18	0.6	309023	3	AE003671 AE003671 Drosophil
c	917	18	0.6	310092	3	AE003826 AE003826 Drosophil
c	918	18	0.6	318200	1	MLEPRTN6 AL583922 Mycobacte
	919	18	0.6	323450	1	SME591790 AL591790 Sinorhizo
c	920	18	0.6	327773	1	AP002554 AP002554 Escherich
	921	18	0.6	329100	1	SME591787 AL591787 Sinorhizo
	922	18	0.6	329861	1	NMA5Z2491 AL162756 Neisseria
	923	18	0.6	349980	6	AX044032 AX044032 Sequence
c	924	18	0.6	349980	6	AX344551 AX344551 Sequence
c	925	17	0.5	54	6	AX019625 AX019625 Sequence
c	926	17	0.5	54	6	AX019626 AX019626 Sequence
	927	17	0.5	77	6	AR104674 AR104674 Sequence
	928	17	0.5	77	6	I38257 I38257 Sequence 87
c	929	17	0.5	78	8	TOMTRALTBM L24059 Lycopersico
	930	17	0.5	106	9	HSCALMI1 U16851 Homo sapien
c	931	17	0.5	107	6	AX019620 AX019620 Sequence
	932	17	0.5	117	6	AR168401 AR168401 Sequence
c	933	17	0.5	117	9	HSZ99552 Z99552 Homo sapien
c	934	17	0.5	120	6	AX019627 AX019627 Sequence
c	935	17	0.5	126	6	AX435716 AX435716 Sequence
c	936	17	0.5	157	9	HSAOX1S20 AF009460 Homo sapi
c	937	17	0.5	170	5	STU60219 U60219 Salmo trutt
c	938	17	0.5	176	6	AX019585 AX019585 Sequence
c	939	17	0.5	214	8	AY038211 AY038211 Sinodiels
c	940	17	0.5	216	8	AF073563S1 AF073563 Naufraga
	941	17	0.5	224	8	AY020225 AY020225 Oryza sat

```
  942    17   0.5    224  8  AY022437          AY022437 Oryza sat
  943    17   0.5    224  8  AY023270          AY023270 Oryza sat
c 944    17   0.5    224  8  AY023482          AY023482 Oryza sat
  945    17   0.5    224  8  AY023920          AY023920 Oryza sat
  946    17   0.5    227  8  AY022893          AY022893 Oryza sat
c 947    17   0.5    227  8  AY023147          AY023147 Oryza sat
  948    17   0.5    228  8  AY024248          AY024248 Oryza sat
  949    17   0.5    230  8  AY023295          AY023295 Oryza sat
  950    17   0.5    280  4  AB054437          AB054437 Megaptera
  951    17   0.5    281  4  AB054492          AB054492 Physeter
  952    17   0.5    281  4  AB054503          AB054503 Berardius
  953    17   0.5    281  10  S34320S2          S34321 Evx-2=even-
  954    17   0.5    282  4  AB054404          AB054404 Balaenopt
  955    17   0.5    283  4  AB054415          AB054415 Platanist
  956    17   0.5    283  4  AB054481          AB054481 Balaenopt
c 957    17   0.5    283  8  RICYK36           D29723 Rice mRNA,
  958    17   0.5    287  9  HSCACNLA35        U30700 Human dihyd
  959    17   0.5    289  6  AX305829          AX305829 Sequence
c 960    17   0.5    291  9  HS18F10F          Z60123 H.sapiens C
  961    17   0.5    303  11  AU048363          AU048363 Rattus no
  962    17   0.5    304  11  G10687            G10687 human STS C
  963    17   0.5    307  9  HS200YB12         Z23510 H. sapiens
  964    17   0.5    307  11  G60422            G60422 SHGC-150076
  965    17   0.5    312  10  AF074961          AF074961 Rattus no
  966    17   0.5    315  6  AX315052          AX315052 Sequence
c 967    17   0.5    317  6  AX184716          AX184716 Sequence
c 968    17   0.5    319  11  G09539            G09539 human STS C
  969    17   0.5    322  6  AR168402          AR168402 Sequence
  970    17   0.5    335  6  AX410988          AX410988 Sequence
  971    17   0.5    335  11  G13047            G13047 sWSS2957 Er
c 972    17   0.5    336  6  AX439176          AX439176 Sequence
  973    17   0.5    352  6  AX071283          AX071283 Sequence
c 974    17   0.5    363  11  G22122            G22122 human STS W
c 975    17   0.5    364  11  HSA222XB9         Z52567 H.sapiens (
  976    17   0.5    389  3  AF344862          AF344862 Heteroder
  977    17   0.5    395  9  AF217527S1        AF217527 Homo sapi
  978    17   0.5    421  6  AX380968          AX380968 Sequence
  979    17   0.5    427  11  G01024            G01024 fruit fly S
  980    17   0.5    429  3  AF377991          AF377991 Anthonomu
  981    17   0.5    429  6  AR128791          AR128791 Sequence
c 982    17   0.5    430  11  G47149            G47149 Z15725_1 Ze
  983    17   0.5    439  6  AX462383          AX462383 Sequence
c 984    17   0.5    440  6  AX321808          AX321808 Sequence
  985    17   0.5    453  4  AF132106          AF132106 Mustela v
  986    17   0.5    455  9  S97393            S97393 growth horm
  987    17   0.5    460  9  AF190059          AF190059 Homo sapi
  988    17   0.5    471  9  AF188896          AF188896 Homo sapi
  989    17   0.5    474  9  AB037480          AB037480 Pongo pyg
  990    17   0.5    475  9  AF188897          AF188897 Homo sapi
  991    17   0.5    478  9  AF189213          AF189213 Homo sapi
c 992    17   0.5    480  6  AX207390          AX207390 Sequence
c 993    17   0.5    489  3  AB054364          AB054364 Batillari
  994    17   0.5    489  9  AF187552          AF187552 Homo sapi
  995    17   0.5    495  9  AF332559          AF332559 Homo sapi
c 996    17   0.5    498  5  AF002619          AF002619 Rivulus h
  997    17   0.5    500  6  AR138493          AR138493 Sequence
  998    17   0.5    501  4  AB054470          AB054470 Pontopori
  999    17   0.5    504  4  AB054371          AB054371 Inia geof
 1000    17   0.5    507  9  AB037475          AB037475 Homo sapi
```

ALIGNMENTS

```
RESULT 1
AX024064
LOCUS       AX024064                3204 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 7 from Patent FR2785293.
ACCESSION   AX024064
VERSION     AX024064.1  GI:10184376
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 3204)
  AUTHORS   Nassif,X., Tinsley,C., Aujame,L., Perrin,A., Rokbi,B.,
            Bouchardon,A. and Renauld,M.G.
  JOURNAL   Patent: FR 2785293-A 7 05-MAY-2000;
            PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES             Location/Qualifiers
     source          1..3204
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
     CDS             1..3204
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC08908.1"
                     /db_xref="GI:10184377"
                     /translation="MRTTPTFPTKTFKPAAMALAVATTLSACLGGGGGTSAPDFNAGG
                     TGIGSNSRATTAKSAAVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPPNLHT
                     GDFTNPNDAYKNLINLKPAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNE
                     NYKNYTAYMRKEAPEDGGGKDIKASFDDEAVIETEAKPTDIRHVKEIGHIDVVSHIIG
                     GRSVDGRPAGGIAPDATLHIMNTHDGTKNEIMSAAIRNAWVKLGERGVRIVNNSFGTT
                     SRAGTADHFQIANSEEQYRQALLAYSGGDKTDEGIRLMQQSDYGNLSYHIRNKNMLFI
                     FSASNDAQAQPNTLTLLPFYEKDAQKGIITVAGVDRSGEKFNGSNHCGITAMWCLSAP
                     YEASVRFTRTNPIQIAGTSFSAPIVTGTAALLLQKYPWMSNDNLRTTLLTTAQDIGAV
                     GVDSKFGWGLLDAGKAMNGPASFPFGDFTADTKGTSDIAYSFRNDISGTGGLIKKGGS
                     QLQLHGNNTYTGKTIIEGGSLVLYGNNKSDMRVETKGALIYNGAASGGSLNSDGIVYL
                     ADTDRSGANETVHIKGDLQLGGEGTLYTRLGKLLKVDGTAMTGGKLYMSARGKGAGYL
                     NRTGQRVPFLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGDTLSYYVRRGNA
                     ARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAADRTDMPGI
                     RPYGATFRAAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLKAVSDGLDHN
                     ATGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHSTWSE
                     NSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTLM
                     QLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLTEGTLVGLAG
                     LKLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGL
                     GADVEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRF"
BASE COUNT      831 a    909 c    899 g    565 t
ORIGIN

  Query Match            100.0%;  Score 3204;  DB 6;  Length 3204;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120

Qy        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180

Qy        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240

Qy        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300

Qy        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360

Qy        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420

Qy        421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480

Qy    481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540

Qy    541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600

Qy    601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660

Qy    661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720

Qy    721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780

Qy    781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840

Qy    841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900

Qy    901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960

Qy    961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020

Qy   1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080

Qy   1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
```

```
Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180

Qy   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||
Db   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
```

```
RESULT 2
NME311654
LOCUS       NME311654               3204 bp    DNA     linear   BCT 24-JUL-2002
DEFINITION  Neisseria meningitidis ausP gene for autotransporter serine
            protease.
ACCESSION   AJ311654
VERSION     AJ311654.1  GI:13397949
KEYWORDS    ausP gene; autotransporter serine protease.
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1
  AUTHORS   Turner,D.P., Wooldridge,K.G. and Ala'Aldeen,D.A.
  TITLE     Autotransported Serine Protease A of Neisseria meningitidis: an
            Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein
  JOURNAL   Infect. Immun. 70 (8), 4447-4461 (2002)
  MEDLINE   22112893
REFERENCE   2  (bases 1 to 3204)
  AUTHORS   Ala'Aldeen,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2001) Ala'Aldeen D.A., Microbiology, University
            of Nottingham, University Hospital, Nottingham, NG7 2UH, UNITED
            KINGDOM
COMMENT     Related sequence: AJ277537.
FEATURES             Location/Qualifiers
     source          1..3204
                     /organism="Neisseria meningitidis"
                     /strain="Group C ET37"
                     /isolate="Z4181"
                     /db_xref="taxon:487"
                     /country="Mali"
     gene            1..3204
                     /gene="ausP"
     CDS             1..3204
                     /gene="ausP"
                     /function="unknown"
                     /codon_start=1
                     /transl_table=11
                     /product="autotransporter serine protease (AusP)"
                     /protein_id="CAC34605.1"
                     /db_xref="GI:13397950"
                     /db_xref="SPTREMBL:Q9AE78"
                     /translation="MRTTPTFPTKTFKPAAMALAVATTLSACLGGGGGTSAPDFNAGG
                     TGIGSNSRATTAKSAAVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPPNLHT
                     GDFTNPNDAYKNLINLKPAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNE
                     NYKNYTAYMRKEAPEDGGGKDIKASFDDEAVIETEAKPTDIRHVKEIGHIDVVSHIIG
                     GRSVDGRPAGGIAPDATLHIMNTHDGTKNEIMSAAIRNAWVKLGERGVRIVNNSFGTT
                     SRAGTADHFQIANSEEQYRQALLAYSGGDKTDEGIRLMQQSDYGNLSYHIRNKNMLFI
                     FSASNDAQAQPNTLALLPFYEKDAQKGIITVAGVDRSGEKFNGSNHCGITAMWCLSAP
                     YEASVRFTRTNPIQIAGTSFSAPIVTGTAALLLQKYPWMSNDNLRTTLLTTAQDIGAV
                     GVDSKFGWGLLDAGKAMNGPASFPFGDFTADTKGTSDIAYSFRNDISGTGGLIKKGGS
                     QLQLHGNNTYTGKTIIEGGSLVLYGNNKSDMRVETKGALIYNGAASGGSLNSDGIVYL
                     ADTDRSGANETVHIKGDLQLGGEGTLYTRLGKLLKVDGTAMTGGKLYMSARGKGAGYL
                     NRTGQRVPFLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGDTLSYYVRRGNA
                     ARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAADRTDMPGI
                     RPYGATFRAAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLKAVSDGLDHN
                     ATGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHSTWSE
                     NSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTLM
                     QLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLTEGTLVGLAG
                     LKLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGL
                     GADVEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRF"
BASE COUNT      830 a     909 c     900 g     565 t
ORIGIN

  Query Match             98.4%;  Score 3153;  DB 1;  Length 3204;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3203;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy     61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120

Qy    121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180

Qy    181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240

Qy    241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300

Qy    301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360

Qy    361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420

Qy    421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAGAACACGGCTATAACGAA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAGAACACGGCTATAACGAA 480

Qy    481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540
```

```
Qy    541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600

Qy    601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660

Qy    661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720

Qy    721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780

Qy    781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840

Qy    841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900

Qy    901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960

Qy    961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020

Qy   1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   1021 GCACAAGCTCAGCCCAACACACTGGCCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080

Qy   1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180

Qy   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||
Db   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204


RESULT 3
NMA2Z2491/c
LOCUS       NMA2Z2491             349061 bp    DNA     linear   BCT 04-DEC-2000
DEFINITION  Neisseria meningitidis serogroup A strain Z2491 complete genome;
            segment 2/7.
ACCESSION   AL162753 AL157959
VERSION     AL162753.2  GI:7379120
KEYWORDS    .
SOURCE      Neisseria meningitidis Z2491.
  ORGANISM  Neisseria meningitidis Z2491
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 349061)
  AUTHORS   Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
            Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
            Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
            Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
            Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
  TITLE     Complete DNA sequence of a serogroup A strain of Neisseria
            meningitidis Z2491
  JOURNAL   Nature 404 (6777), 502-506 (2000)
  MEDLINE   20222556
   PUBMED   10761919
REFERENCE   2  (bases 1 to 349061)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
            sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of N. meningitidis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES             Location/Qualifiers
     source          1. .349061
                     /organism="Neisseria meningitidis Z2491"
                     /strain="Z2491"
                     /db_xref="taxon:122587"
                     /note="serogroup: A"
     repeat_unit     complement(24. .206)
                     /note="ATR repeat; hmmfs hit to HMM ATR (1 - 183), score:
                     310.39"
                     /label=ATR
     RBS             209. .212
     gene            220. .681
                     /gene="NMA0368"
     CDS             220. .681
                     /gene="NMA0368"
                     /note="NMA0368, probable integral membrane protein, len:
                     153 aa; contains four probable transmembrane domains"
                     /codon_start=1
                     /transl_table=11
                     /product="putative integral membrane protein"
                     /protein_id="CAB83669.1"
                     /db_xref="GI:7379121"
                     /db_xref="SPTREMBL:Q9JWH7"
                     /translation="MQEQNRKSSFPIVMLLVSVALWIASLSNVAFYLGNHGSMEGLTV
                     LILGSIFASLDIRYCAVYANYVWLAAIVLLALRKKVVPVHAAFWGLALVAFSVKAVYV
                     DEAGNTSDIVRYGAGFYLWYAAFAVASIGTFAGKNKERKAASAADGTKNDV"
     gene            671. .1492
                     /gene="hemK"
     CDS             671. .1492
                     /gene="hemK"
                     /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
                     to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
                     scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
                     PS00092 N-6 Adenine-specific DNA methylases signature"
                     /codon_start=1
                     /transl_table=11
                     /product="HemK protein"
                     /protein_id="CAB83670.1"
                     /db_xref="GI:7379122"
                     /db_xref="SPTREMBL:Q9JWH6"
                     /translation="MTFDKWLGLSKLPKNEARMLLQYVSEYTRVQLLTRGGEEMPDEV
                     RQRADRLAQRRLNGEPVAYILGAREFYGRRFTVNPSVLIPRPETEHLVEAVLARLPEN
                     GRVWDLGTGSGAVAVTVALERPDAFVRASDISPPALETARKNAADLGARVEFAYGSWF
                     DTDMPSEGKWDIIVSNPPYIENGDKHLSQGDLRFEPQIALTDFSDGLSCIRTLAQGAP
                     DRLAEGGFLLLEHGFDQGAAVRGVLAENGFSGVETLPDLAGLDRVTLGKYMKHLK"
     misc_feature    831. .840
                     /gene="hemK"
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     misc_feature    1162. .1171
                     /gene="hemK"
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     misc_feature    1187. .1207
                     /gene="hemK"
                     /note="PS00092 N-6 Adenine-specific DNA methylases
                     signature"
     misc_feature    complement(1279. .1288)
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     RBS             1565. .1568
     gene            1578. .2966
                     /gene="NMA0370"
     CDS             1578. .2966
                     /gene="NMA0370"
                     /note="NMA0370, probable integral membrane protein, len:
                     462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
                     P44640 hypothetical protein HI0325 (450 aa), fasta scores;
                     E(): 0, 49.8% identity in 464 aa overlap"
                     /codon_start=1
                     /transl_table=11
                     /product="putative integral membrane protein"
                     /protein_id="CAB83671.1"
                     /db_xref="GI:7379123"
                     /db_xref="SPTREMBL:Q9JWH5"
                     /translation="MNAVVVAVIVMLVLSLSRVHVVLSLTIGAFVGGAVAGMPLQNIA
                     DAAGQVSQAGIIPVFNKGLEGGAKIALSYAMLGAFAMAITHSGLPQQLAGAVVRKLNR
```

```
                GGMPDSVRSGEGVVKWLLLSIILVMGMMSQNIIPIHIAFIPMIVPPLLLVFNRLKIDR
                RLIACVITFGLVTTYMFLPYGFGAIFLNEILLGNIHSAAPQLDVKNINVMAAMAIPAL
                GMLAGLLLAFVHYRKPRLYQSNNADTAGNADAANRPQPSAYRSLAAAVAIAVCFAIQL
                MYEDSLVLGAMLGFAVFMMLGVINRDKANDVFGEGIKMMAMVGFIMIAAQGFAAVMNA
                TGHIQPLVESSMAIFGNSKGMAALAMLVVGLLVTMGIGSSFSTLPIIAAIYVPLCVGL
                GFSPLATVAIVGTAGALGDAGSPASDSTLGPTMGLNADGQHDHIRDSVIPTFIHYNIP
                LLIAGWIAAMVL"
gene            2967..3191
                /gene="slyX"
CDS             2967..3191
                /gene="slyX"
                /note="NMA0371, slyX, SLYX protein homolog, len: 74 aa;
                simlar to SLYX_HAEIN P44759 SLYX protein homolog (73 aa),
                fasta scores; E(): 0.0017, 33.8% identity in 74 aa
                overlap, and SLYX_ECOLI P30857 SLYX protein (72 aa), fasta
                scores; E(): 0.24, 32.4% identity in 68 aa overlap"
                /codon_start=1
                /transl_table=11
                /product="SLYX protein homolog"
                /protein_id="CAB83672.1"
                /db_xref="GI:7379124"
                /db_xref="SPTREMBL:Q9JWH4"
                /translation="MDAVQEFEHRITELEIQSALQEDVIAGLNAMVAELRQTLDLQQA
                QLRLLYQKMQDRNPDAQEPYSLRDEIPPHY"
gene            complement(3271..3423)
                /gene="NMA0372"
CDS             complement(3271..3423)
                /gene="NMA0372"
                /note="NMA0372, unknown, questionable CDS, len: 50 aa"
                /codon_start=1
                /transl_table=11
                /product="very hypothetical protein NMA0372"
                /protein_id="CAB83673.1"
                /db_xref="GI:7379125"
                /db_xref="SPTREMBL:Q9JWH3"
                /translation="MFGRLSPLGSDGFFLAVYEYPVGFICLQGKTAYFRSGGKRFHRI
                KGILSD"
misc_feature    3562..3571
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
gene            complement(3586..4356)
                /gene="thiF"
CDS             complement(3586..4356)
                /gene="thiF"
                /note="NMA0373, thiF, probable ThiF protein, len: 256 aa;
                similar to e.g. THIF_ECOLI P30138 THIF protein (251 aa),
                fasta scores; E(): 0, 43.1% identity in 246 aa overlap,
                and MOEB_ECOLI P12282 molybdopterin biosynthesis MOEB
                protein. (249 aa), fasta scores; E(): 0, 43.9% identity in
                244 aa overlap (note that N.m. does not have orthologs of
                any other molybdopterin biosynthesis proteins). Contains
                Pfam match to entry PF00899 ThiF_family, ThiF family"
                /codon_start=1
                /transl_table=11
                /product="ThiF protein"
                /protein_id="CAB83674.1"
                /db_xref="GI:7379126"
                /db_xref="SPTREMBL:Q9JWH2"
                /translation="MTTTEHDNDDAFLLRYSRHILLDEIGIEGQQKLSAAHILVVGCG
                GLGAAALPYLAASGIGTLTIADSDTVELHNLQRQVAFDEGDVGKLKTEALADRLRHIN
                HTVDVRTINEKLDGCRLTGLVQAADIVLDCCDNYATRQAVNRACVQAKTPLVSGAAVR
                FEGQLAVYRPDLPDSPCYACLFDGGSASDGICSLFGVFSPLVGIIGSTQAAEALKILL
                DAGEPSHGRLAVYRALEGGWQYFDLPRNPECPVCGAER"
misc_feature    3787..3796
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    complement(3847..4257)
                /gene="thiF"
                /note="Pfam match to entry PF00899 ThiF_family, ThiF
                family, score 186.60, E-value 4.1e-52"
misc_feature    complement(4418..4427)
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
gene            4471..7224
                /gene="ppc"
CDS             4471..7224
                /gene="ppc"
                /EC_number="4.1.1.31"
                /note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
                917 aa; simlar to many e.g. CAPP_RHOPA O32483
                phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),
                fasta scores; E(): 0, 43.3% identity in 928 aa overlap.
                Contains 2x Pfam match to entry PF00311 PEPcase,
                Phosphoenolpyruvate carboxylase, PS00017 ATP/GTP-binding
                site motif A (P-loop), and PS00393 Phosphoenolpyruvate
                carboxylase active site 2"
                /codon_start=1
                /transl_table=11
                /product="phosphoenolpyruvate carboxylase"
                /protein_id="CAB83675.1"
                /db_xref="GI:7379127"

  Query Match             67.3%;  Score 2156;  DB 1;  Length 349061;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 3006;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy        182 TATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTC 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126136 TATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTC 126077

Qy        242 GGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCGAATCTGC 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126076 GGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCGAATCTGC 126017

Qy        302 ATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTG 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126016 ATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTG 125957

Qy        362 CAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAAT 421
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     125956 CAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAAT 125897

Qy        422 CCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125896 CCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAA 125837

Qy        482 ATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAG 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     125836 ATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGACGGTAAAG 125777

Qy        542 ACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATA 601
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125776 ACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATA 125717

Qy        602 TCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTT 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125716 TCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTT 125657

Qy        662 CCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATA 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125656 CCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATA 125597

Qy        722 CGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGC 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125596 CGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGC 125537

Qy        782 TGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCA 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125536 TGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCA 125477

Qy        842 CTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCT 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125476 CTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCT 125417

Qy        902 ATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCA 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 125416 ATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCA 125357

Qy    962 ACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACG 1021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125356 ACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACG 125297

Qy   1022 CACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAG 1081
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db 125296 CACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAG 125237

Qy   1082 GCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATT 1141
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db 125236 GCATTATCACAGTTGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATT 125177

Qy   1142 GCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACCC 1201
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 125176 GCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCC 125117

Qy   1202 GTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGG 1261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125116 GTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGG 125057

Qy   1262 CGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGC 1321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125056 CGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGC 124997

Qy   1322 TGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGC 1381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124996 TGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGC 124937

Qy   1382 TGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCG 1441
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124936 TGGATACGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCG 124877

Qy   1442 ATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCG 1501
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124876 ATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCG 124817

Qy   1502 GCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCA 1561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124816 GCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCA 124757

Qy   1562 AAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCG 1621
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db 124756 AAACCGTTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCACG 124697

Qy   1622 TCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCG 1681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124696 TCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCG 124637

Qy   1682 ACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATCA 1741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124636 ACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATCA 124577

Qy   1742 AAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGA 1801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124576 AAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGA 124517

Qy   1802 AAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGGG 1861
          ||||||||||||||||||||||||||||||||||||||||||||||| |||| || ||||
Db 124516 AAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCAGGCGGTAAGGGGG 124457

Qy   1862 CAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGC 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124456 CAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGC 124397

Qy   1922 GGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGACA 1981
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db 124396 GGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGTCTGCTGGCTTCCCTCGACA 124337

Qy   1982 GCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCA 2041
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 124336 GCGTCGAAAAAACAGCGGGTAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCA 124277

Qy   2042 ATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCG 2101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124276 ATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCG 124217

Qy   2102 TAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCAT 2161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124216 TAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCAT 124157

Qy   2162 CCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCC 2221
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124156 CCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCC 124097

Qy   2222 GCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACG 2281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124096 GCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACG 124037

Qy   2282 GTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATG 2341
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db 124036 GTGTACGCATCTTCAACAATCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATG 123977

Qy   2342 CCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGG 2401
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 123976 CCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGG 123917

Qy   2402 GTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTG 2461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123916 GTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTG 123857

Qy   2462 AAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATA 2521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123856 AAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATA 123797

Qy   2522 CGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATG 2581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123796 CGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATG 123737

Qy   2582 CAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCT 2641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123736 CAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCT 123677

Qy   2642 ATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTG 2701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123676 ATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTG 123617

Qy   2702 CGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCG 2761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123616 CGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCG 123557

Qy   2762 GTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACG 2821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123556 GTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACG 123497

Qy   2822 ACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACA 2881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123496 ACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACA 123437

Qy   2882 GCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCG 2941
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123436 GCATCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCG 123377

Qy   2942 ATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACA 3001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123376 ATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACA 123317

Qy   3002 CGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATA 3061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123316 CGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATA 123257

Qy   3062 TGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGA 3121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123256 TGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGA 123197
```

```
Qy    3122 ACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAG 3181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123196 ACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAG 123137

Qy    3182 TCGGCGTAGGCTACCGGTTCTGA 3204
           |||||||||||||||||||||||
Db 123136 TCGGCGTAGGCTACCGGTTCTGA 123114
```

```
RESULT 4
AX236469
LOCUS       AX236469                 4218 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 162 from Patent WO0164922.
ACCESSION   AX236469
VERSION     AX236469.1  GI:15796056
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4218)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 162 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4218
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961cL-983"
BASE COUNT     1196 a   1149 c   1130 g    743 t
ORIGIN

  Query Match             26.0%;  Score 834;  DB 6;  Length 4218;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2971 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3030

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3031 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3090

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3091 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3150

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    3151 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3210

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3211 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3270

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3271 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3330

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    3331 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3390

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           ||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db    3391 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3450

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3451 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3510

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    3511 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3570

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3571 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3630

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3631 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3690

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3691 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3750

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3751 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3810

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3811 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3870

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    3871 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3930

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3931 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3990

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3991 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4050

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    4051 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4110

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4111 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4170

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
           ||||||||||||||||||||||||||||||||||||||||||
Db    4171 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212
```

```
RESULT 5
AX239755
LOCUS       AX239755                 4218 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 57 from Patent WO0164920.
ACCESSION   AX239755
VERSION     AX239755.1  GI:15797388
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4218)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 57 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4218
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961cL-983"
BASE COUNT     1196 a   1149 c   1130 g    743 t
ORIGIN
```

```
  Query Match             26.0%;  Score 834;  DB 6;  Length 4218;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2971 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3030

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3031 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3090

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3091 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3150

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       3151 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3210

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3211 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3270

Qy       2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3271 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3330

Qy       2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       3331 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3390

Qy       2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db       3391 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3450

Qy       2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3451 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3510

Qy       2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3511 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3570

Qy       2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3571 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3630

Qy       2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3631 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3690

Qy       2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3691 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3750

Qy       2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3751 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3810

Qy       2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3811 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3870

Qy       2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       3871 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3930

Qy       2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3931 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3990

Qy       2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3991 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4050

Qy       3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
              |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db       4051 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4110

Qy       3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4111 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4170

Qy       3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
              ||||||||||||||||||||||||||||||||||||||||||
Db       4171 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212


RESULT 6
AE002545
LOCUS       AE002545               10591 bp    DNA     linear   BCT 25-MAY-2000
DEFINITION  Neisseria meningitidis serogroup B strain MC58 section 187 of 206
            of the complete genome.
ACCESSION   AE002545 AE002098
VERSION     AE002545.1  GI:7227229
KEYWORDS    .
SOURCE      Neisseria meningitidis MC58.
  ORGANISM  Neisseria meningitidis MC58
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 10591)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Complete genome sequence of Neisseria meningitidis serogroup B
            strain MC58
  JOURNAL   Science 287 (5459), 1809-1815 (2000)
  MEDLINE   20175755
   PUBMED   10710307
REFERENCE   2  (bases 1 to 10591)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..10591
                     /organism="Neisseria meningitidis MC58"
                     /strain="MC58"
                     /db_xref="taxon:122586"
                     /note="serogroup: B"
     gene            548..3796
                     /gene="NMB1969"
     CDS             548..3796
                     /gene="NMB1969"
                     /note="similar to SP:P31631 GB:U07788 GB:M62363 PID:507431
                     percent identity: 48.64; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="serotype-1-specific antigen, putative"
                     /protein_id="AAF42298.1"
```

```
                     /db_xref="GI:7227230"
                     /translation="MRTTPTFPTKTFKPTAMALAVATTLSACLGGGGGGTSAPDFNAG
                     GTGIGSNSRATTAKSAAVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPPNLH
                     TGDFPNPNDAYKNLINLKPAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYN
                     ENYKNYTAYMRKEAPEDGGGKDIEASFDDEAVIETEAKPTDIRHVKEIGHIDLVSHII
                     GGRSVDGRPAGGIAPDATLHIMNTNDETKNEMMVAAIRNAWVKLGERGVRIVNNSFGT
                     TSRAGTADLFQIANSEEQYRQALLDYSGGDKTDEGIRLMQQSDYGNLSYHIRNKNMLF
                     IFSTGNDAQAQPNTYALLPFYEKDAQKGIITVAGVDRSGEKFKREMYGEPGTEPLEYG
                     SNHCGITAMWCLSAPYEASVRFTRTNPIQIAGTSFSAPIVTGTAALLLQKYPWMSNDN
                     LRTTLLTTAQDIGAVGVDSKFGWGLLDAGKAMNGPASFPFGDFTADTKGTSDIAYSFR
                     NDISGTGGLIKKGGSQLQLHGNNTYTGKTIIEGGSLVLYGNNKSDMRVETKGALIYNG
                     AASGGSLNSDGIVYLADTDQSGANETVHIKGSLQLDGKGTLYTRLGKLLKVDGTAIIG
                     GKLYMSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAG
                     SEGDTLSYYVRRGNAARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPE
                     TVETAAADRTDMPGIRPYGATFRAAAAVQHANAADGVRIFNSLAATVYADSTAAHADM
                     QGRRLKAVSDGLDHNGTGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENT
                     TAAATLGMGRSTWSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRST
                     GADEHAEGSVNGTLMQLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGW
                     SGNSLTEGTLVGLAGLKLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGK
                     TGARNMPHTRLVAGLGADVEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRF"
     gene            4828..6609
                     /gene="NMB1970"
     CDS             4828..6609
                     /gene="NMB1970"
                     /note="similar to GB:AE000511 PID:2313385 percent
                     identity: 53.22; identified by sequence similarity;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="para-aminobenzoate synthetase component
                     I/4-amino-4-deoxychorismate lyase, putative"
                     /protein_id="AAF42299.1"
                     /db_xref="GI:7227231"
                     /translation="MPYFALFDDAVSGRAKRYQNHVESRFFRPEELDALDGALQSGWQ
                     KGLHSVLFADYGFGLPLTGVESERGGNLALHWFANCADIDAESWLARHSDGLPAGIST
                     PQPSVSETDYLDRIRQIHEAIRRGDTYQINYTTRLHLQAYGNPVSLYRRLRQPVPYAV
                     LSHLPDAEGQSAWTLCFSPELFLKIGSDGTISTEPMKGTAPILGDGQDERRAAELQAD
                     PKNRAENVMIVDLLRNDLGKIAQTGTVCVPEPFKVSRFGSVWQMTSTIQAQALPHTSF
                     ADILRAAFPCGSITGAPKKMSMQIIESLEAEARGLYTGSIGYLNPCSGGLGFEGTFNV
                     VIRTLSLTPLSDGIYQGVYGVGSGIVIDSDPAAEYRECGWKARFLNELRPDFGIFETL
                     RAENGRCTLLDRHLCRLKTSAQALNLPLPDGCENQIKQYIADLPDGAFRVKALLASDG
                     ISLSRAVLNRLTDKQRVIISPAVLPAQNYLRRFKTTCRALFDQAWQTAETQGAFDSLF
                     FNSDGILLEGGRSNVFIKHRGQWLTPSLDLDILNGIMRQAVLDEPQKYLQTNQVIETH
                     ITQKTLQEAEEIRLSNALRGVFAAALA"
     gene            complement(6668..8185)
                     /gene="NMB1971"
     CDS             complement(6668..8185)
                     /gene="NMB1971"
                     /note="conserved hypothetical protein; identified by
                     Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF42300.1"
                     /db_xref="GI:7227232"
                     /translation="MLYFRYGFLVVWCAAGVSAAYGADAPAILDDKALLQVQRSVSDK
                     WAESDWKVENDAPRVVDGDFLLAHPKMLEHSLRDALNGNQADLIASLADLYAKLPDYD
                     AVLYGRARALLAKLAGRPAEAVARYRELHGENAADERILLDLAAAEFDDFRLKSAERH
                     FAEAAKLDLPAPVLENVGRFRKKTEGLTGWRFSGGISPAVNRNANNAAPQYCRQNGGR
                     QICSVSRAERAAGLNYEIEAEKLTPLADNHYLLFRSNIGGTSYYFSKKSAYDDGFGRA
                     YLGWQYKNARQTAGILPFYQVQLSGSDGFDAKTKRVNNRRLPPYMLAHGVGVQLSHTY
                     RPNPGWQFSVALEHYRQRYREQDRAEYNNGRQDGFYVSSAKRLGESATVFGGWQFVRF
                     VPKRETVGGAVNNAAYRRNGVYAGWAQEWRQLGGLNSRVSASYARRNYKGIAAFSTEA
                     QRNREWNVSLALSHDKLSYKGIVPALNYRFGRTESNVPYAKRRNSEVFVSADWRF"
     gene            complement(8881..10515)
                     /gene="NMB1972"
     CDS             complement(8881..10515)
                     /gene="NMB1972"
                     /note="similar to SP:P29842 GB:Z23008 GB:Z23009 PID:438185
                     PID:438187 percent identity: 99.26; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
```

```
                     /product="chaperonin, 60 kDa"
                     /protein_id="AAF42301.1"
                     /db_xref="GI:7227233"
                     /translation="MAAKDVQFGNEVRQKMVNGVNILANAVRVTLGPKGRNVVVDRAF
                     GGPHITKDGVTVAKEIELKDKFENMGAQMVKEVASKTNDVAGDGTTTATVLAQSIVAE
                     GMKYVTAGMNPTDLKRGIDKAVAALVDELKNIAKPCDTSKEIAQVGSISANSDEQVGA
                     IIAEAMEKVGKEGVITVEDGKSLENELDVVEGMQFDRGYLSPYFINDAEKQIAALDNP
                     FVLLFDKKISNIRDLLPVLEQVAKASRPLLIIAEDVEGEALATLVVNNIRGILKTVAV
                     KAPGFGDRRKAMLQDIAILTGGVVISEEVGLSLEKATLDDLGQAKRIEIGKENTTIID
                     GFGDAAQIEARVAEIRQQIETATSDYDKEKLQERVAKLAGGVAVIKVGAATEVEMKEK
                     KDRVEDALHATRAAVEEGVVAGGGVALLRARAALENLHTGNADQDAGVQIVLRAVESP
                     LRQIVANAGGEPSVVVNKVLEGKGNYGYNAGSGEYGDMIEMGVLDPAKVTRSALQHAA
                     SIAGLMLTTDCMIAEIPEDKPAVPDMGGMGGMGGMM"
BASE COUNT     2623 a   3083 c   2609 g   2276 t
ORIGIN

  Query Match             26.0%;  Score 834;  DB 1;  Length 10591;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2555 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2614

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2615 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2674

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2734

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2735 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2794

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2795 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2854

Qy       2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2855 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2914

Qy       2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       2915 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2974

Qy       2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db       2975 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3034

Qy       2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3035 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3094

Qy       2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3095 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3154

Qy       2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3155 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3214

Qy       2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3215 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3274

Qy       2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3275 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3334

Qy       2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3335 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3394

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3395 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3454

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     3455 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3514

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3515 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3574

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3575 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3634

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     3635 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3694

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3695 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3754

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
            ||||||||||||||||||||||||||||||||||||||||||
Db     3755 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3796


RESULT 7
AX044034
LOCUS       AX044034               349980 bp    DNA     linear   PAT 24-NOV-2000
DEFINITION  Sequence 113 from Patent WO0066791.
ACCESSION   AX044034
VERSION     AX044034.1  GI:11342918
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 349980)
  AUTHORS   Pizza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
            Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
            Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
  TITLE     Neisseria genomic sequences and methods of their use
  JOURNAL   Patent: WO 0066791-A 113 09-NOV-2000;
            CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1..349980
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
                     /note="sequence too long, cut in 8 pieces.~seq 1: 1 to
                     349980 349980 bases~seq 108: 300001 to 649980 349980
                     bases~seq 109: 600001 to 949980 349980 bases~seq 110:
                     900001 to 1249980 349980 bases~seq 111: 1200001 to 1549980
                     349980 bases~seq 112: 1500001 to 1849980 349980 bases~seq
                     113: 1800001 to 2149980 349980 bases~seq 114: 2100001 to
                     2272325 172325 bases"
BASE COUNT    86473 a  95646 c  85908 g  81953 t
ORIGIN

  Query Match             26.0%;  Score 834;  DB 6;  Length 349980;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263419 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 263478

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263479 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 263538

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263539 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 263598

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   263599 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 263658

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263659 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 263718

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263719 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 263778

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   263779 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 263838

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   263839 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 263898

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263899 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 263958

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACACAGCACATGG 2562
            |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   263959 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACGCAGCACATGG 264018

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264019 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 264078

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264079 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 264138

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264139 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 264198

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264199 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 264258

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264259 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 264318

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   264319 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 264378

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264379 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 264438

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264439 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 264498

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   264499 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 264558

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 264559 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 264618

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||||||||||||||||||||
Db 264619 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 264660


RESULT 8
AX236409
LOCUS       AX236409                3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 102 from Patent WO0164922.
ACCESSION   AX236409
VERSION     AX236409.1  GI:15796027
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 102 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1. .3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-741"
BASE COUNT     1046 a   1100 c   1107 g    686 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144


RESULT 9
AX236419
LOCUS       AX236419                3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 112 from Patent WO0164922.
ACCESSION   AX236419
VERSION     AX236419.1  GI:15796032
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 112 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1. .3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG741-983"
BASE COUNT     1046 a   1103 c   1106 g    684 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
```

```
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2674 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2733

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2734 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2793

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2794 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2853

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2854 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2913

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2914 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2973

Qy       2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2974 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3033

Qy       2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       3034 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3093

Qy       2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db       3094 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3153

Qy       2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3154 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3213

Qy       2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3214 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3273

Qy       2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3274 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3333

Qy       2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3334 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3393

Qy       2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3394 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3453

Qy       2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3454 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3513

Qy       2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3514 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3573

Qy       2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       3574 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3633

Qy       2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3634 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3693

Qy       2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3694 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3753

Qy       3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
              |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db       3754 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3813

Qy       3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3814 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3873

Qy       3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
              |||||||||||||||||||||||||||||||||||||||
Db       3874 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912


RESULT 10
AX239717
LOCUS       AX239717               3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 19 from Patent WO0164920.
ACCESSION   AX239717
VERSION     AX239717.1  GI:15797370
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 19 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-741"
BASE COUNT     1046 a   1100 c   1107 g    686 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy       2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy       2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy       2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db       2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385
```

```
Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||||||||||||||||||||
Db   3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144


RESULT 11
AX239729
LOCUS       AX239729                3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 31 from Patent WO0164920.
ACCESSION   AX239729
VERSION     AX239729.1  GI:15797375
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 31 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG741-983"
BASE COUNT     1046 a   1103 c   1106 g    684 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2674 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2733

Qy   2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2734 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2793

Qy   2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2794 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2853

Qy   2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   2854 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2913

Qy   2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2914 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2973

Qy   2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2974 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3033

Qy   2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   3034 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3093

Qy   2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
          |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   3094 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3153

Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3154 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3213

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3214 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3273

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3274 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3333

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3334 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3393

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3394 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3453

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3454 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3513

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3514 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3573

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   3574 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3633
```

```
Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3634 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3693

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3694 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3753

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     3754 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3813

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3814 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3873

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     3874 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912


RESULT 12
AX236463
LOCUS       AX236463                4170 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 156 from Patent WO0164922.
ACCESSION   AX236463
VERSION     AX236463.1  GI:15796053
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4170)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 156 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4170
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961c-983"
BASE COUNT     1185 a   1138 c   1119 g    728 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4170;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2905 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2964

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2965 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3024

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3025 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3084

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     3085 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3144

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3145 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3204

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3205 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3264

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     3265 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3324

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     3325 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3384

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3385 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3444

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     3445 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3504

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3505 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3564

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3565 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3624

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3625 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3684

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3685 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3744

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3745 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3804

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     3805 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3864

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3865 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3924

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3925 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3984

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     3985 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4044

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4045 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4104

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     4105 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143


RESULT 13
AX239749
LOCUS       AX239749                4170 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 51 from Patent WO0164920.
ACCESSION   AX239749
VERSION     AX239749.1  GI:15797385
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4170)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
```

```
                and Pizza,M.
  TITLE         Hybrid expression of neisserial proteins
  JOURNAL       Patent: WO 0164920-A 51 07-SEP-2001;
                Chiron Spa (IT)
FEATURES                 Location/Qualifiers
     source              1..4170
                         /organism="synthetic construct"
                         /db_xref="taxon:32630"
                         /note="961c-983"
BASE COUNT     1185 a   1138 c   1119 g    728 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4170;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2905 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2964

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2965 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3024

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3025 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3084

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       3085 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3144

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3145 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3204

Qy       2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3205 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3264

Qy       2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       3265 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3324

Qy       2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db       3325 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3384

Qy       2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3385 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3444

Qy       2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3445 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3504

Qy       2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3505 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3564

Qy       2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3565 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3624

Qy       2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3625 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3684

Qy       2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3685 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3744

Qy       2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3745 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3804

Qy       2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       3805 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3864

Qy       2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3865 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3924

Qy       2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3925 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3984

Qy       3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
              ||||||||||||||||||||||||||||||| |||||||||||| |||||||||||||||
Db       3985 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4044

Qy       3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4045 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4104

Qy       3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
              |||||||||||||||||||||||||||||||||||||||
Db       4105 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143


RESULT 14
AX236413
LOCUS       AX236413                4179 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 106 from Patent WO0164922.
ACCESSION   AX236413
VERSION     AX236413.1  GI:15796029
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4179)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 106 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4179
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961c"
BASE COUNT     1187 a   1140 c   1121 g    731 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4179;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy      2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy      2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
             |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db      2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy      2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
             |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db      2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy      2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy      2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy      2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy      2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy      2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy      2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy      2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy      2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy      2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy      2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy      3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
             |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db      2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy      3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy      3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
             |||||||||||||||||||||||||||||||||||||||
Db      3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 15
AX239721
LOCUS       AX239721                4179 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 23 from Patent WO0164920.
ACCESSION   AX239721
VERSION     AX239721.1  GI:15797372
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4179)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 23 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4179
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961c"
BASE COUNT     1187 a   1140 c   1121 g    731 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4179;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy      2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy      2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy      2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy      2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy      2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy      2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
             |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db      2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy      2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
             |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db      2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy      2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy      2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy      2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy      2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy      2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||||||||||||||||||||
Db   3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 16
AX236457
LOCUS       AX236457                4335 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 150 from Patent WO0164922.
ACCESSION   AX236457
VERSION     AX236457.1  GI:15796050
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4335)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 150 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961-983"
BASE COUNT     1217 a   1189 c   1165 g    764 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4335;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3070 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3129

Qy   2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3130 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3189

Qy   2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3190 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3249

Qy   2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   3250 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3309

Qy   2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3310 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3369

Qy   2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3370 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3429

Qy   2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   3430 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3489

Qy   2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
          |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   3490 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3549

Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3550 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3609

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3610 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3669

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3670 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3729

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3730 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3789

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3790 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3849

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3850 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3909

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3910 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3969

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   3970 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 4029

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4030 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 4089

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4090 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4149

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   4150 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4209

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4210 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4269

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||||||||||||||||||||
Db   4270 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308
```

```
RESULT 17
AX239743
LOCUS       AX239743                4335 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 45 from Patent WO0164920.
ACCESSION   AX239743
VERSION     AX239743.1  GI:15797382
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4335)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 45 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961-983"
BASE COUNT     1217 a   1189 c   1165 g    764 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4335;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3070 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3129

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3130 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3189

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3190 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3249

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    3250 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3309

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3310 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3369

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3370 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3429

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    3430 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3489

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    3490 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3549

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3550 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3609

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    3610 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3669

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3670 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3729

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3730 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3789

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3790 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3849

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3850 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3909

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3910 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3969

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    3970 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 4029

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4030 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 4089

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4090 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4149

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    4150 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4209

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4210 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4269

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||||||||||||||||||||
Db    4270 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308


RESULT 18
AX236411
LOCUS       AX236411                4344 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 104 from Patent WO0164922.
ACCESSION   AX236411
VERSION     AX236411.1  GI:15796028
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4344)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 104 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4344
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961"
BASE COUNT     1219 a   1191 c   1167 g    767 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4344;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965
```

```
Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144


RESULT 19
AX239719
LOCUS       AX239719                4344 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 21 from Patent WO0164920.
ACCESSION   AX239719
VERSION     AX239719.1  GI:15797371
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4344)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 21 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4344
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961"
BASE COUNT     1219 a   1191 c   1167 g    767 t
ORIGIN

  Query Match             25.9%;  Score 831;  DB 6;  Length 4344;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445
```

```
Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||||||||||||||||||||
Db    3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 20
AX236407
LOCUS       AX236407              4425 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 100 from Patent WO0164922.
ACCESSION   AX236407
VERSION     AX236407.1  GI:15796026
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4425)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 100 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1. .4425
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-ORF46.1"
BASE COUNT     1165 a   1259 c   1219 g    782 t
ORIGIN
```

```
  Query Match             25.9%;  Score 831;  DB 6;  Length 4425;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| ||||||||||||||
Db     2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 21
AX239715
LOCUS       AX239715                4425 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 17 from Patent WO0164920.
ACCESSION   AX239715
VERSION     AX239715.1  GI:15797369
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4425)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 17 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4425
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-ORF46.1"
BASE COUNT     1165 a   1259 c   1219 g    782 t
ORIGIN

  Query Match            25.9%;  Score 831;  DB 6;  Length 4425;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| ||||||||||||||
Db     2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 22
NME277537
LOCUS       NME277537               3254 bp    DNA     linear   BCT 24-JUL-2002
DEFINITION  Neisseria meningitidis ausP gene for autotransporter serine
            protease.
ACCESSION   AJ277537
VERSION     AJ277537.2  GI:13539237
KEYWORDS    ausP gene; autotransporter serine protease.
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1
  AUTHORS   Turner,D.P., Wooldridge,K.G. and Ala'Aldeen,D.A.
  TITLE     Autotransported Serine Protease A of Neisseria meningitidis: an
            Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein
  JOURNAL   Infect. Immun. 70 (8), 4447-4461 (2002)
```

```
  MEDLINE   22112893
REFERENCE   2
  AUTHORS   Ala'Aldeen,D.A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Ala'Aldeen D.A.A., Microbiology, University
            of Nottingham, Meningococcal Research Group, Division of
            Microbiology, University Hospital, Nottingham NG7 2UH, UNITED
            KINGDOM
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 3254)
  AUTHORS   Ala'Aldeen,D.A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2001) Ala'Aldeen D.A.A., Microbiology, University
            of Nottingham, Meningococcal Research Group, Division of
            Microbiology, University Hospital, Nottingham NG7 2UH, UNITED
            KINGDOM
COMMENT     On Apr 3, 2001 this sequence version replaced gi:7649689.
FEATURES             Location/Qualifiers
     source          1..3254
                     /organism="Neisseria meningitidis"
                     /strain="B:15:P1.16"
                     /isolate="SD"
                     /db_xref="taxon:487"
                     /country="United Kingdom:England"
     gene            1..3254
                     /gene="ausP"
                     /pseudo
     CDS             1..3254
                     /gene="ausP"
                     /function="unknown"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="autotransporter serine protease"
BASE COUNT      845 a    922 c    917 g    570 t
ORIGIN

  Query Match           24.4%;  Score 783;  DB 1;  Length 3254;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1233;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2013 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2072

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2073 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2132

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2133 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2192

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    2193 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2252

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2253 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2312

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2313 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2372

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    2373 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2432

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    2433 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2492

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    2493 TGGGAACAGGGCGGTGTTGAGGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2552

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACACAGCACATGG 2562
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    2553 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACGCAGCACATGG 2612

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2613 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2672

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2673 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2732

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2733 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2792

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2793 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2852

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2853 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2912

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    2913 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2972

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2973 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3032

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3033 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3092

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    3093 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3152

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3153 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3212

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
           ||||||||||||||||||||||||||||||||||||||||||
Db    3213 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3254


RESULT 23
AF169448/c
LOCUS       AF169448                 600 bp    DNA     linear   BCT 09-AUG-2000
DEFINITION  Neisseria meningitidis strain Z2491 clone Cm024 unknown sequence.
ACCESSION   AF169448
VERSION     AF169448.1  GI:9754656
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Identification of regions of the chromosome of Neisseria
            meningitidis and Neisseria gonorrhoeae which are specific to
            pathogenic Neisseriae
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 600)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
            rue de Vaugirard, Paris 75015, France
FEATURES             Location/Qualifiers
     source          1..600
                     /organism="Neisseria meningitidis"
                     /strain="Z2491"
                     /db_xref="taxon:487"
                     /clone="Cm024"
BASE COUNT      111 a    159 c    144 g    184 t      2 others
ORIGIN

  Query Match             10.0%;  Score 321;  DB 1;  Length 600;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-168;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        565 GAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGA 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGA 282

Qy        625 CACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGC 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 CACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGC 222

Qy        685 GGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAA 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAA 162

Qy        745 ATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATC 804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATC 102

Qy        805 GTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCC 864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCC 42

Qy        865 AATTCGGAGGAGCAGTACCGC 885
              |||||||||||||||||||||
Db         41 AATTCGGAGGAGCAGTACCGC 21


RESULT 24
AF169473
LOCUS       AF169473               286 bp    DNA     linear   BCT 09-AUG-2000
DEFINITION  Neisseria meningitidis strain Z2491 clone Em085 unknown sequence.
ACCESSION   AF169473
VERSION     AF169473.1  GI:9754681
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 286)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Identification of regions of the chromosome of Neisseria
            meningitidis and Neisseria gonorrhoeae which are specific to
            pathogenic Neisseriae
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 286)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
            rue de Vaugirard, Paris 75015, France
FEATURES             Location/Qualifiers
     source          1..286
                     /organism="Neisseria meningitidis"
                     /strain="Z2491"
                     /db_xref="taxon:487"
                     /clone="Em085"
BASE COUNT       88 a     73 c     63 g     62 t
ORIGIN

  Query Match              5.5%;  Score 177;  DB 1;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-87;
  Matches  177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy        925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy        985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 1041
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 177


RESULT 25
A68924
LOCUS       A68924                 286 bp    DNA     linear   PAT 06-MAY-1999
DEFINITION  Sequence 95 from Patent WO9802547.
ACCESSION   A68924
VERSION     A68924.1  GI:4759843
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 286)
  AUTHORS   Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
            Merker,P.
  TITLE     DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
            SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
            APPLICATIONS
  JOURNAL   Patent: WO 9802547-A 95 22-JAN-1998;
            INST NAT SANTE RECH MED (FR)
COMMENT     Other publication FR 2751000 19980116.
FEATURES             Location/Qualifiers
     source          1..286
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       88 a     73 c     63 g     62 t
ORIGIN

  Query Match              5.5%;  Score 177;  DB 6;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-87;
  Matches  177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy        925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy        985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 1041
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 177


RESULT 26
AX236815
LOCUS       AX236815                50 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 508 from Patent WO0164922.
ACCESSION   AX236815
VERSION     AX236815.1  GI:15796401
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 508 07-SEP-2001;
```

```
          Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..50
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide"
BASE COUNT       15 a     20 c      7 g      8 t
ORIGIN

  Query Match             1.1%;  Score 35;  DB 6;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-07;
  Matches   35;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 CGAACGACCCCAACCTTCCCTACAAAAACTTTCAA 38
              |||||||||||||||||||||||||||||||||||
Db       16 CGAACGACCCCAACCTTCCCTACAAAAACTTTCAA 50


RESULT 27
NGOPAK
LOCUS       NGOPAK                  1434 bp    DNA     linear   BCT 10-APR-1995
DEFINITION  N.gonorrhoeae opaK gene for opacity protein.
ACCESSION   X52364 S36068 S75784
VERSION     X52364.1  GI:48702
KEYWORDS    opacity protein; opaK gene.
SOURCE      Neisseria gonorrhoeae.
  ORGANISM  Neisseria gonorrhoeae
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1434)
  AUTHORS   Bhat,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-1990) Bhat K., National Institutes of Health,
            Rocky Mountain Laboratory, Hamilton 59840, U S A
REFERENCE   2  (bases 1 to 1318)
  AUTHORS   Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
            Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
  TITLE     The opacity proteins of Neisseria gonorrhoeae strain MS11 are
            encoded by a family of 11 complete genes
  JOURNAL   Mol. Microbiol. 5 (8), 1889-1901 (1991)
  MEDLINE   92114767
   PUBMED   1815562
REFERENCE   3  (bases 1 to 1434)
  AUTHORS   Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
            Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
  TITLE     The opacity proteins of Neisseria gonorrhoeae strain MS11 are
            encoded by a family of 11 complete genes
  JOURNAL   Mol. Microbiol. 6 (8), 1073-1076 (1992)
  MEDLINE   92261323
   PUBMED   1584024
COMMENT     The cloned gene is in its non-expressed state. The change in number
            of the coding repeat unit has caused the atg start codon in the
            signal peptide to be out of frame with the rest of its coding
            sequence.
FEATURES             Location/Qualifiers
     source          1..1434
                     /organism="Neisseria gonorrhoeae"
                     /strain="MS11"
                     /db_xref="taxon:485"
     repeat_unit     218..243
     repeat_unit     344..369
     -35_signal      373..380
     -10_signal      398..403
     RBS             455..460
     gene            465..1255
                     /gene="opaK"
     CDS             465..1255
                     /gene="opaK"
                     /note="The cloned gene is in its non-expressed state. The
                     change in the number of coding repeat unit has caused the
                     atg start codon to be out of frame with the rest of its
                     coding sequence"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="opacity protein"
                     /db_xref="PID:e145998"
     repeat_region   490..524
                     /note="signal peptide coding sequence"
                     /rpt_type=TANDEM
     mat_peptide     539..1252
                     /gene="opaK"
                     /product="opacity protein"
     terminator      1269..1294
BASE COUNT      383 a    425 c    343 g    283 t
ORIGIN

  Query Match             1.0%;  Score 33;  DB 1;  Length 1434;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-06;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     3146 CCGGTTCCAAACAGTACGGCAACCACAGCGGAC 3178
              |||||||||||||||||||||||||||||||||
Db       30 CCGGTTCCAAACAGTACGGCAACCACAGCGGAC 62


RESULT 28
AC017356/c
LOCUS       AC017356               37497 bp    DNA     linear   HTG 09-DEC-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC017356
VERSION     AC017356.1  GI:6553630
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 37497)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10210411 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..37497
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT    10676 a   8030 c   7947 g  10844 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 2;  Length 37497;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db    26162 GCAGCAACAGCAGAGCAACAACAGC 26138


RESULT 29
AC009749/c
LOCUS       AC009749              161278 bp    DNA     linear   INV 28-JUL-2001
DEFINITION  Drosophila melanogaster, chromosome 2L, region 25B-25C, BAC clone
            BACR40I04, complete sequence.
ACCESSION   AC009749
VERSION     AC009749.8  GI:15027726
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
```

```
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 161278)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 2L, region 25B-25C
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 161278)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Jul 28, 2001 this sequence version replaced gi:6980162.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .161278
                     /organism="Drosophila melanogaster"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="2L"
                     /map="25B-25C"
                     /clone="BACR40I04 (D1088)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
BASE COUNT    44503 a  35245 c  35792 g  45738 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 3;  Length 161278;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db  119005 GCAGCAACAGCAGAGCAACAACAGC 118981


RESULT 30
AE003575
LOCUS       AE003575            327209 bp    DNA     linear   INV 04-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386046 section 16
            of 16, complete sequence.
ACCESSION   AE003575 AE002638
VERSION     AE003575.2  GI:10727324
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 327209)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
   PUBMED   10731132
REFERENCE   2  (bases 1 to 327209)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     On Oct 9, 2000 this sequence version replaced gi:7295650.
FEATURES             Location/Qualifiers
     source          1. .327209
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="2L"
     gene            <29815. .>30955
                     /gene="CG11933"
                     /evidence=not_experimental
                     /map="25A3-25A3"
                     /db_xref="FLYBASE:FBan0011933"
                     /db_xref="FLYBASE:FBgn0031615"
     mRNA            join(<29815. .29907,30773. .>30955)
                     /gene="CG11933"
                     /product="CT35796"
                     /evidence=not_experimental
                     /db_xref="FLYBASE:FBan0011933"
                     /db_xref="FLYBASE:FBgn0031615"
     CDS             join(29815. .29907,30773. .30955)
```

```
                     /gene="CG11933"
                     /note="CG11933 gene product"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF50997.1"
                     /db_xref="GI:7295691"
                     /db_xref="FLYBASE:FBan0011933"
                     /db_xref="FLYBASE:FBgn0031615"
                     /translation="MAQPFLGSPDCHLWTGELGGSQGFCGSSSGCDYLRHHLSPNFQD
                     YWDQRARVHREYQIFPGQMPLQVGVHLDLRFLPTNCQEYNSNQAQVQ"
     gene            complement(<31022. .>31300)
                     /gene="CG11932"
                     /evidence=not_experimental
                     /map="25A3-25A3"
                     /db_xref="FLYBASE:FBan0011932"
                     /db_xref="FLYBASE:FBgn0031616"
     mRNA            complement(<31022. .>31300)
                     /gene="CG11932"
                     /product="CT35794"
                     /evidence=not_experimental
                     /db_xref="FLYBASE:FBan0011932"
                     /db_xref="FLYBASE:FBgn0031616"
     CDS             complement(31022. .31300)
                     /gene="CG11932"
                     /note="CG11932 gene product"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF50996.1"
                     /db_xref="GI:7295690"
                     /db_xref="FLYBASE:FBan0011932"
                     /db_xref="FLYBASE:FBgn0031616"
                     /translation="MVEVGRKDSKIYANSKLEVWAFPEDRQDDWQTPIDSNLVVPVVL
                     IGEREEKMILVDSNLVVWAMPEDGRDDQQIPMDSDLVFPVVLVVKGEE"
     gene            <32109. .>35468
                     /gene="CG15635"
                     /evidence=not_experimental
                     /map="25A3-25A3"
                     /db_xref="FLYBASE:FBan0015635"
                     /db_xref="FLYBASE:FBgn0031617"
     mRNA            join(<32109. .33305,33412. .34156,34267. .>35468)
                     /gene="CG15635"
                     /product="CT35795"
                     /evidence=not_experimental
                     /db_xref="FLYBASE:FBan0015635"
                     /db_xref="FLYBASE:FBgn0031617"
     CDS             join(32109. .33305,33412. .34156,34267. .35468)
                     /gene="CG15635"
                     /note="CG15635 gene product"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF50995.2"
                     /db_xref="GI:10727325"
                     /db_xref="FLYBASE:FBan0015635"
                     /db_xref="FLYBASE:FBgn0031617"
                     /translation="MEFSFDRIRSGESYTPFHFPYVPDNFPTPSAPAPPPKPRPPPPP
                     PPPPTTTRPPTTTPTPTTTPTPITTPPPPPPSAPPPPDPATPGVWYPLPIYGNAPQFG
                     DPPGSHLLSNSGKPHPPGSQMPPNSGFVLPPAICSPPYGPPNQDGNPPGSQTPPNSEL
                     VPPPAIWNPAYEPPNQIGHPPESQAPTNSGFVPPPAIWNPPYGPPNPDGNPPESQTPP
                     NSGFVPPPAIWSPSVGPTSEDGHPPESQAPSNSGFVPPLAICYPPFGQQNQIGHQPGS
                     HKPPNSGIYPPSTGWIPPSGSLSQGGHPPGSQLPPNSGLPPGSIPPLRPPNQIGDPPG
                     SQMPSNLGIYPPSMGWIPPSGSLIQGGHPPGSLYPPNSVLAPGSIPPLRPSNQGGHPP
                     GSLLPPNRALFRLLDHQIKLDNHQDHCCLLMWEYIPHPRAGFPPSGPLTQGGHPTGSL
                     LPPNTGLPPGSIPPLRPPNQGGHPPGSQKPPNVGIYPPSTGWIPPFGPLTQGGHPPGS
                     LLPPNTGLPPGSIPPLGSPNQIGHPPGSQKPPNVGIYPPSTGWTPPSGPLTQGGHPPG
                     SLLPPNTGLPPGSIPPLGSPNQIGHPPGSQKPPNVGIYPPSTGWIPPSGPLTQSGHPP
                     GSLLPPNTGLPPGSIPPLRPPNQGSLLPPNTGLPSGSIPPLGSPNQIGHPPGSQKPPN
                     SGIYPPSTGSIPPSGPLTQGGHPPGSLLPSNTGLPSGSIPPLGSPNQIGHPPGSQKPP
                     NSGIYPPSTGWIPPSGPLTQGGHPPGSLLPPNTGLPPGSISPLGSPNQIGHPPGSQKP
                     PNVGIYPPSTGSIPPLRPPNQGGHPPGSQKPPNVGIYPPSTGWIPPSGPLTQGGHPPG
                     SQVPSNSVLPPGSIPPLRPPNQGRHPPGSQVPPNTGLPPGSIPPLGSPNQIGHPPGSQ
                     KPPNSGIYPPSTGSIPPSGPLTQGGHPPGSLLPPNTGLPPGSIPPLRPPNQGRHPPGS
                     QVPPNSVLPPGSIPPLGSPIQIGRPVGSQKPASSKIRPPPPSEPQNPDNVPKQPVKNP
                     PAGGNLNVNDIRIDRPN"
     gene            <36096. .>36760
                     /gene="CG11931"
                     /evidence=not_experimental
                     /map="25A3-25A3"
                     /db_xref="FLYBASE:FBan0011931"
                     /db_xref="FLYBASE:FBgn0031618"
     mRNA            join(<36096. .36173,36602. .>36760)
                     /gene="CG11931"
                     /product="CT35793"
                     /evidence=not_experimental
                     /db_xref="FLYBASE:FBan0011931"
                     /db_xref="FLYBASE:FBgn0031618"
     CDS             join(36096. .36173,36602. .36760)
                     /gene="CG11931"
                     /note="CG11931 gene product"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF50994.1"
                     /db_xref="GI:7295688"
                     /db_xref="FLYBASE:FBan0011931"
                     /db_xref="FLYBASE:FBgn0031618"
                     /translation="MCKWQHIFILISLVPHGHGLERRDSSPYYYSYYNYYYSYYYYYY
                     YYYYYYYSYYYYYYYYYSYYSCYYYYSYYYYFS"
     gene            <40399. .>41772
                     /gene="CG3355"
                     /map="25A3-25A3"
                     /db_xref="FLYBASE:FBan0003355"
                     /db_xref="FLYBASE:FBgn0031619"
     mRNA            join(40399. .40684,40761. .41152,41419. .>41772)
                     /gene="CG3355"
                     /product="CT11187"
                     /db_xref="FLYBASE:FBan0003355"
                     /db_xref="FLYBASE:FBgn0031619"
     CDS             join(40486. .40684,40761. .41152,41419. .41772)
                     /gene="CG3355"
                     /note="CG3355 gene product"
                     /codon_start=1
                     /protein_id="AAF50993.1"
                     /db_xref="GI:7295687"
                     /db_xref="FLYBASE:FBan0003355"
                     /db_xref="FLYBASE:FBgn0031619"
                     /translation="MRVYLALPLLLLGIGLSLAQYQYQAPHQTLAQQFADVVDVVDPA
                     EQSIKAVRPPKSRNQCTAKQNCFCGTPNVNRIVGGQQVRSNKYPWTAQLVKGRHYPRL
                     FCGGSLINDRYVLTAAHCVHGNRDQITIRLLQIDRSSRDPGIVRKVVQTTVHPNYDPN
                     RIVNDVALLKLESPVPLTGNMRPVCLPEANHNFDGKTAVVAGWGLIKEGGVTSNYLQE
                     VNVPVITNAQCRQTRYKDKIAEVMLCAGLVQQGGKDACQGDSGGPLIVNEGRYKLAGV
                     VSFGYGCAQKNAPGVYARVSKFLDWIRKNTADGCYCQS"
     gene            complement(<44894. .>45106)
                     /gene="CG11930"

  Query Match             0.8%;  Score 25;  DB 3;  Length 327209;
  Best Local Similarity   100.0%;  Pred. No. 0.16;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db  280767 GCAGCAACAGCAGAGCAACAACAGC 280791


RESULT 31
AX024142/c
LOCUS       AX024142                  30 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 85 from Patent FR2785293.
ACCESSION   AX024142
VERSION     AX024142.1  GI:10184453
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Nassif,X., Tinsley,C., Aujame,L., Perrin,A., Rokbi,B.,
            Bouchardon,A. and Renauld,M.G.
  JOURNAL   Patent: FR 2785293-A 85 05-MAY-2000;
```

```
            PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="amorce de"
BASE COUNT        7 a     11 c      9 g      3 t
ORIGIN

  Query Match              0.7%;  Score 23;  DB 6;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   3181 GTCGGCGTAGGCTACCGGTTCTG 3203
          |||||||||||||||||||||||
Db     30 GTCGGCGTAGGCTACCGGTTCTG 8


RESULT 32
AB014075/c
LOCUS       AB014075               14043 bp    DNA     linear   BCT 08-MAY-1999
DEFINITION  Clostridium histolyticum genes for hypoxanthine-guanine
            phosphoribosyl-transferase (HGPRTase), GTPase and 12 ORFs, complete
            and partial cds.
ACCESSION   AB014075
VERSION     AB014075.1  GI:3868863
KEYWORDS    ruvB; colH; Orf2u; Orf3u; Orf4u; Orf5u; Orf6u; hflX; GTPase; hprT;
            hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase); Orf9u.
SOURCE      Clostridium histolyticum (strain:JCM 1403) vegitative cell DNA.
  ORGANISM  Clostridium histolyticum
            Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
            Clostridium.
REFERENCE   1  (sites)
  AUTHORS   Matsushita,O., Jung,C.M., Katayama,S., Minami,J., Takahashi,Y. and
            Okabe,A.
  TITLE     Gene duplication and multiplicity of collagenases in Clostridium
            histolyticum
  JOURNAL   J. Bacteriol. 181 (3), 923-933 (1999)
  MEDLINE   99121032
REFERENCE   2  (bases 1 to 14043)
  AUTHORS   Matsushita,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAY-1998) Osamu Matsushita, Kagawa Medical
            University, Department of Microbiology; 1750-1 Ikenobe, Miki-cho,
            Kagawa 761-0793, Japan (E-mail:osamu@kms.ac.jp,
            Tel:+81(87)891-2129, Fax:+81(87)891-2129)
FEATURES             Location/Qualifiers
     source          1. .14043
                     /organism="Clostridium histolyticum"
                     /strain="JCM 1403"
                     /db_xref="taxon:1498"
                     /dev_stage="vegitative cell"
     CDS             153. .800
                     /note="Orf9u"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34252.1"
                     /db_xref="GI:3868864"
                     /translation="MKFNIKAKLISLIICIIIIITFVYLFDSIVTPTVLAVSEAEMR
                     AKTIEIINGCIIDEVEKGFNYNDIIKIDKDDEGNISLIRADILKLNKIASNISLKGQK
                     KLNELGRIGIKLSLGYITKNNILSYIGPDIAVKMEPTSNVETTYSSIFKSAGINQTSH
                     KIFIKFKTKIRIIIPLKSVDVEVYSEMPISETIIVGKIPDSAIQLDLNKAGYKNN"
     gene            complement(818. .1345)
                     /gene="hprT"
     CDS             complement(818. .1345)
                     /gene="hprT"
                     /codon_start=1
                     /transl_table=11
                     /product="hypoxanthine-guanine phosphoribosyl-transferase
                     (HGPRTase)"
                     /protein_id="BAA34253.1"
                     /db_xref="GI:3868865"
                     /translation="MEDKKRNILIGKEEIASKIQEIGAIISKDYEGKNLYVLSLLRGS
                     FIFTADLVRELQVKTKIGFMTTSSYGNSENSSGTVKVVNDVPDNIEGFDVLIVDDIMD
                     TGITMEFVVKHLKELGANSVNTCVLLDKPERRVAKIKPDYTCFEIPDLFVVGYGLNYG
                     DYYRNIPYIFNWENN"
     gene            1469. .3259
                     /gene="hflX"
     CDS             1469. .3259
                     /gene="hflX"
                     /codon_start=1
                     /transl_table=11
                     /product="GTPase"
                     /protein_id="BAA34254.1"
                     /db_xref="GI:3868866"
                     /translation="MIYGNIEGLKDSILERLEKIYKMPVWKDSIFTEEMVSVIEEVTL
                     KINREISIAISRKGKVVSVAVGDSSTVDLPLIDIKENRLSGVRIIHTHPNGNPKLSAI
                     DISALTKLKLDCIVAIGVGEENRRKVNVGYCSIENDTLCCDEIGPLTIEETISLRFLE
                     RVEFVEKLIKQSEIYEEQGERAVLVGIESEESLEELEELAEACNVETVCSIFQRRERA
                     DNAFLVGSGKVDEISMALQAYKGNVVIFDEELTGAQVRNLEENLGTKVIDRTTLILEI
                     FASRARSREAKLQVELAQLRYRSARLIGLGTILSRTGGGIGTRGPGEKKLEIDKRRIR
                     ERITELSRELKKITSIREIQRDKRVRGNIPKISLVGYTNTGKSTLRNKLCELTPVKEN
                     INKEKVFEANMLFATLDTTTRAIALPDGRACTLSDTVGFIRKLPHDLVESFKSTLEEV
                     VNSNLLLHVVDSSSEKLYEQIDAVEMVLSQLGVDPKECILVLNKIDKLTKEELEDIEL
                     KLNKYKLVTISAKTGENLDLILNAICENLPQTLREITFLIPYTEQGKASFIHRNAKVL
                     DENYEDDGTKIKALVDDEVYNKCMEYEVKC"
     CDS             3272. .4168
                     /note="Orf6u"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34255.1"
                     /db_xref="GI:3868867"
                     /translation="MPKTISEYQEAFNEALENLKTKDSVAAIVVFGSIVNGDIWEKSD
                     IDMMVISSEKQEKIVNIYSDYDGIPIHVRLIHKDELMNLCSGDNKGSHIHRLFSSSKL
                     VYSKDTEIDRLYNEERFYSDIHREKWNVTFLSKLLKSVSECRKYLYMDSLYNAYTLCN
                     IVLRDYSKLVVNSSGYIINRDIVTMAISTDEELSREVEQLFNDKDNLEKCIGRFIRFI
                     EGKINENIRIYASIIIQYLRESDNALSSSEIRDKLGFKHYDINIEEILDSIYELNIIN
                     RCKRGYKTLNEETLIDENVYYL"
     CDS             4180. .5580
                     /note="Orf5u"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34256.1"
                     /db_xref="GI:3868868"
                     /translation="MEIYYIEFKLDKPKYVQISQKIKVIIERGIINDGEKLPSIRKLS
                     ELLNVNKDTVINAYRKLEEEGYAIQKIGSGTYAKKRDTGSRFKKDYSNTIKNLSYKNQ
                     EKYIDLVGEATISSYFPIKNFKTVLNEVLDRDGSKAFIYKEVLGYEGLRKSINNFFWD
                     GKRKSEEILIVSGAQQGIDVISKAIINVRDNIIVEKPTYNGALSVFKWRRANIIEVPI
                     ENDGIDLDQFEKILQKNKIKCFYAMSYFQNPTGVSYSMKKKRILELASIYDFYIIED
                     DYLSELIYDESITYNSFKSLDINDRVIYIKSFSKIFLPGIRLGYLICPKKFLDSTINC
                     KINTDITTSSLMQRALDLYINRGYWKEHIDMLRNNYKETYNVLTEVLAKELSEEVSFN
                     LPKGGLNLYLKLHNTNSIDFFYKCKEQNIITTPGVFFYKNYKEGLDYVKLGFSNLRTD
                     QILSVTNILKEIFKTK"
     CDS             5599. .6273
                     /note="Orf4u"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34257.1"
                     /db_xref="GI:3868869"
                     /translation="MKKGLINFAWRDIDKFSSEEISYYLFLEGKSIDVISKIRSISRE
                     EIQSHIINGKIKFGILSKASNEKELFEYLCTVAKEDRKRSLGFLSVNLKAELLNYIIY
                     NYNSMDWSKKERAIWIIGEAKAENAISILLKASVHKNINLRRLAISALGKLENLKGES
                     ALIRALEDDNPQVLQYAIKSLQKINSTRGIEKIKKIYENTDKKYLKDTCETYLQFVED
                     IMRENI"
     CDS             6286. .6936
                     /note="Orf3u"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34258.1"
                     /db_xref="GI:3868870"
                     /translation="MSYFTIKSKANSEFQEKKSIFIGHVKRVNSEEEAKQFIEEIKEK
                     YKDARHNVYAYIIGENMGIQRYSDDGEPKGTGGIPILEVIKKNNITDTVVVVTRYFGG
                     ILLGASGLTRAYSKAAVNAIKEAGVVEKIIGAKVNIDVEYELLGKIQYTFSQKNIYIE
                     DIQYSDKVTITVFLELDKIADVEREVMNITSGNGILLRDDAKYFFKLEDRLYEEEQ"
     CDS             7024. .7764
                     /note="Orf2u"
```

```
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34259.1"
                     /db_xref="GI:3868871"
                     /translation="MSGHSKWHNIQAKKSKVDAKRGKIFTKIGKEIAIAAKNGGPNPD
                     ANPKLRDVIAKAKANNMPKDSIERAIKKAAGELAGVNYEEILYEGYGPDGIAVLVQAL
                     TDNKNRSAGNVRHAFSKHGGNLGSTGCVSFMFQTKGQIVIEKNDELDEEELMMMALEC
                     GAEDFQSEDEVYIITTSPEDFGSVRETLEEKGLEFLEAEVKSIPDTYTVIDENTAGKF
                     QKMLDVLEDDEDVQDVYHNAEFPEGWEE"
     CDS             complement(7932..8411)
                     /note="Orflu"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34541.1"
                     /db_xref="GI:3892645"
                     /translation="MNHKKEKEYKKLDTEKLKENLTELQYNVTQRNATEKPFLNKYDK
                     HFEDGIYVDIVSGEPLFLSIDKFNSGCGWPAFSKPISRKYIKERADFSHGMSRIEVRS
                     KNADSHLGHVFNDGPIENGGMRYCINSASLKFIAKDKLKEEGYEEFLPLFEKDKQEL"
     gene            8715..11780
                     /gene="colH"
     CDS             8715..11780
                     /gene="colH"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34542.1"
                     /db_xref="GI:3892646"
                     /translation="MKRKCLSKRLMLAITMATIFTVNSTLPIYAAVDKNNATAAVQNE
                     SKRYTVSYLKTLNYYDLVDLLVKTEIENLPDLFQYSSDAKEFYGNKTRMSFIMDEIGR
                     RAPQYTEIDHKGIPTLVEVVRAGFYLGFHNKELNEINKRSFKERVIPSILAIQKNPNF
                     KLGTEVQDKIVSATGLLAGNETAPPEVVNNFTPILQDCIKNIDRYALDDLKSKALFNV
                     LAAPTYDITEYLRATKEKPENTPWYGKIDGFINELKKLALYGKINDNNSWIIDNGIYH
                     IAPLGKLHSNNKIGIETLTEVMKVYPYLSMQHLQSADQIKRHYDSKDAEGNKIPLDKF
                     KKEGKEKYCPKTYTFDDGKVIIKAGARVEEEKVKRLYWASKEVNSQFFRVYGIDKPLE
                     EGNPDDILTMVIYNSPEEYKLNSVLYGYDTNNGGMYIEPEGTFFTYEREAQESTYTLE
                     ELFRHEYTHYLQGRYAVPGQWGRTKLYDNDRLTWYEEGGAELFAGSTRTSGILPRKSI
                     VSNIHNTTRNNRYKLSDTVHSKYGASFEFYNYACMFMDYMYNKDMGILNKLNDLAKNN
                     DVDGYDNYIRDLSSNYALNDKYQDHMQERIDNYENLTVPFVADDYLVRHAYKNPNEIY
                     SEISEVAKLKDAKSEVKKSQYFSTFTLRGSYTGGASKGKLEDQKAMNKFIDDSLKKLD
                     TYSWSGYKTLTAYFTNYKVDSSNRVTYDVVFHGYLPNEGDSKNSLPYGKINGTYKGTE
                     KEKIKFSSEGSFDPDGKIVSYEWDFGDGNKSNEENPEHSYDKVGTYTVKLKVTDDKGE
                     SSVSTTTAEIKDLSENKLPVIYMHVPKSGALNQKVVFYGKGTYDPDGSIAGYQWDFGD
                     GSDFSSEQNPSHVYTKKGEYTVTLRVMDSSGQMSEKTMKIKITDPVYPIGTEKEPNNS
                     KETASGPIVPGIPVSGTIENTSDQDYFYFDVITPGEVKIDINKLGYGGATWVVYDENN
                     NAVSYATDDGQNLSGKFKADKPGRYYIHLYMFNGSYMPYRINIEGSVGR"
     gene            11943..12461
                     /gene="dcd"
     CDS             11943..12461
                     /gene="dcd"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34543.1"
                     /db_xref="GI:3892647"
                     /translation="MILSGKEIKNRLNKDIFIEPFSDNQLNPNSYNLRLHNELLVYEN
                     NVLDMKKENKAKKITIPEEGLLLEPGKLYLGRTIEHTRTEKLVPMLEGRSSVGRLGLF
                     IHITAGFGDIGFSGFWTLEIFCVQPIRIYPNIEICQIYYHNIEGEYEKYTSGKYQNNT
                     GVQPSLLFKDFE"
     CDS             12526..13140
                     /note="orf2d"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34544.1"
                     /db_xref="GI:3892648"
                     /translation="MSNKKKLLIFIPSFFIAIFLLTYGISTFENRNHKASSNKQGEAK
                     TGEKINIDNSSNSIKIVYESVYLKPEGEQNKTTVDECIEKFNKDKKVISDMTKEQVLA
                     VFKKHDYNLKDIKKDQIVFSRSINKYKYQEGKYVIGIKDGSVAIFKVDKNGNMLLESP
                     KDITTIKVEQLPEGDIELLQKGDKIYEFNTKEEAVEGVEAIFRT"
     gene            13269..13862
                     /gene="ruvA"
     CDS             13269..13862
                     /gene="ruvA"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA34545.1"
                     /db_xref="GI:3892649"

  Query Match             0.7%;  Score 23;  DB 1;  Length 14043;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        537 TAAAGACATTAAAGCTTCTTTCG 559
              |||||||||||||||||||||||
Db      13760 TAAAGACATTAAAGCTTCTTTCG 13738


RESULT 33
AC013119/c
LOCUS       AC013119               48626 bp    DNA     linear   HTG 03-NOV-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC013119
VERSION     AC013119.1  GI:6223212
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 48626)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10213447 by the submitter.
            For further information on this sequence you may e-mail to
            fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..48626
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT    13541 a  10753 c  10456 g  13876 t
ORIGIN

  Query Match             0.7%;  Score 22;  DB 2;  Length 48626;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 GGGATGCCAAAATCAATGCCCC 287
              ||||||||||||||||||||||
Db      31038 GGGATGCCAAAATCAATGCCCC 31017


RESULT 34
DMBR33H2/c
LOCUS       DMBR33H2               85918 bp    DNA     linear   HTG 11-OCT-1999
DEFINITION  Drosophila melanogaster chromosome X clone BACR33H2 map 4B-C strain
            y; cn bw sp, *** SEQUENCING IN PROGRESS ***, in unordered pieces.
ACCESSION   AL121812
VERSION     AL121812.1  GI:6048199
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 85918)
  AUTHORS   RA Murphy,L., Harris,D. and Barrell,B.
  TITLE     Sequencing the X chromosome of Drosophila melanogaster
  JOURNAL   Unpublished
  REMARK    Sanger Centre, Wellcome Trust Genome Campus, Hinxton Hall, Hinxton,
            Cambridge CB10 1SA, U.K.
REFERENCE   2  (bases 1 to 85918)
  AUTHORS   Benos,P.
```

```
TITLE     Direct Submission
JOURNAL   Submitted (10-OCT-1999) European Drosophila Genome Sequencing
          Consortium
COMMENT   This is a 'working draft' sequence.  It currently
          consists of 147 contigs. The true order of the pieces is not known
          and their order in this sequence record is arbitrary. Gaps between
          the contigs are represented as runs of N, but the exact sizes of
          the gaps are unknown. This record will be updated with the finished
          sequence. 1       62: contig of 62 in length
          63      162: gap of unknown length
          163      624: contig of 462 in length
          625      724: gap of unknown length
          725      846: contig of 122 in length
          847      946: gap of unknown length
          947     1318: contig of 372 in length
          1319     1418: gap of unknown length
          1419     1973: contig of 555 in length
          1974     2073: gap of unknown length
          2074     3024: contig of 951 in length
          3025     3124: gap of unknown length
          3125     3182: contig of 58 in length
          3183     3282: gap of unknown length
          3283     3435: contig of 153 in length
          3436     3535: gap of unknown length
          3536     3908: contig of 373 in length
          3909     4008: gap of unknown length
          4009     4511: contig of 503 in length
          4512     4611: gap of unknown length
          4612     4940: contig of 329 in length
          4941     5040: gap of unknown length
          5041     5355: contig of 315 in length
          5356     5455: gap of unknown length
          5456     5599: contig of 144 in length
          5600     5699: gap of unknown length
          5700     6066: contig of 367 in length
          6067     6166: gap of unknown length
          6167     6207: contig of 41 in length
          6208     6307: gap of unknown length
          6308     6339: contig of 32 in length
          6340     6439: gap of unknown length
          6440     6729: contig of 290 in length
          6730     6829: gap of unknown length
          6830     7282: contig of 453 in length
          7283     7382: gap of unknown length
          7383     7678: contig of 296 in length
          7679     7778: gap of unknown length
          7779     8219: contig of 441 in length
          8220     8319: gap of unknown length
          8320     8768: contig of 449 in length
          8769     8868: gap of unknown length
          8869     8913: contig of 45 in length
          8914     9013: gap of unknown length
          9014     9518: contig of 505 in length
          9519     9618: gap of unknown length
          9619     9649: contig of 31 in length
          9650     9749: gap of unknown length
          9750    10202: contig of 453 in length
          10203     10302: gap of unknown length
          10303     10772: contig of 470 in length
          10773     10872: gap of unknown length
          10873     10915: contig of 43 in length
          10916     11015: gap of unknown length
          11016     11396: contig of 381 in length
          11397     11496: gap of unknown length
          11497     12072: contig of 576 in length
          12073     12172: gap of unknown length
          12173     12536: contig of 364 in length
          12537     12636: gap of unknown length
          12637     13083: contig of 447 in length
          13084     13183: gap of unknown length
          13184     13549: contig of 366 in length
          13550     13649: gap of unknown length
          13650     13957: contig of 308 in length
          13958     14057: gap of unknown length
          14058     14532: contig of 475 in length
          14533     14632: gap of unknown length
          14633     15119: contig of 487 in length
          15120     15219: gap of unknown length
          15220     15725: contig of 506 in length
          15726     15825: gap of unknown length
          15826     16252: contig of 427 in length
          16253     16352: gap of unknown length
          16353     16376: contig of 24 in length
          16377     16476: gap of unknown length
          16477     16816: contig of 340 in length
          16817     16916: gap of unknown length
          16917     17261: contig of 345 in length
          17262     17361: gap of unknown length
          17362     17703: contig of 342 in length
          17704     17803: gap of unknown length
          17804     18078: contig of 275 in length
          18079     18178: gap of unknown length
          18179     19666: contig of 1488 in length
          19667     19766: gap of unknown length
          19767     19857: contig of 91 in length
          19858     19957: gap of unknown length
          19958     19996: contig of 39 in length
          19997     20096: gap of unknown length
          20097     20399: contig of 303 in length
          20400     20499: gap of unknown length
          20500     20940: contig of 441 in length
          20941     21040: gap of unknown length
          21041     21544: contig of 504 in length
          21545     21644: gap of unknown length
          21645     22022: contig of 378 in length
          22023     22122: gap of unknown length
          22123     22429: contig of 307 in length
          22430     22529: gap of unknown length
          22530     23046: contig of 517 in length
          23047     23146: gap of unknown length
          23147     23604: contig of 458 in length
          23605     23704: gap of unknown length
          23705     24181: contig of 477 in length
          24182     24281: gap of unknown length
          24282     24710: contig of 429 in length
          24711     24810: gap of unknown length
          24811     25023: contig of 213 in length
          25024     25123: gap of unknown length
          25124     25580: contig of 457 in length
          25581     25680: gap of unknown length
          25681     25969: contig of 289 in length
          25970     26069: gap of unknown length
          26070     26501: contig of 432 in length
          26502     26601: gap of unknown length
          26602     26908: contig of 307 in length
          26909     27008: gap of unknown length
          27009     27056: contig of 48 in length
          27057     27156: gap of unknown length
          27157     27446: contig of 290 in length
          27447     27546: gap of unknown length
          27547     27740: contig of 194 in length
          27741     27840: gap of unknown length
          27841     28149: contig of 309 in length
          28150     28249: gap of unknown length
          28250     28576: contig of 327 in length
          28577     28676: gap of unknown length
          28677     29126: contig of 450 in length
          29127     29226: gap of unknown length
          29227     29984: contig of 758 in length
          29985     30084: gap of unknown length
          30085     30513: contig of 429 in length
          30514     30613: gap of unknown length
          30614     31009: contig of 396 in length
          31010     31109: gap of unknown length
          31110     31486: contig of 377 in length
          31487     31586: gap of unknown length
```

```
        31587    31864: contig of 278 in length
        31865    31964: gap of unknown length
        31965    32519: contig of 555 in length
        32520    32619: gap of unknown length
        32620    33277: contig of 658 in length
        33278    33377: gap of unknown length
        33378    33681: contig of 304 in length
        33682    33781: gap of unknown length
        33782    34039: contig of 258 in length
        34040    34139: gap of unknown length
        34140    34535: contig of 396 in length
        34536    34635: gap of unknown length
        34636    35081: contig of 446 in length
        35082    35181: gap of unknown length
        35182    35871: contig of 690 in length
        35872    35971: gap of unknown length
        35972    36469: contig of 498 in length
        36470    36569: gap of unknown length
        36570    37081: contig of 512 in length
        37082    37181: gap of unknown length
        37182    37644: contig of 463 in length
        37645    37744: gap of unknown length
        37745    38170: contig of 426 in length
        38171    38270: gap of unknown length
        38271    38821: contig of 551 in length
        38822    38921: gap of unknown length
        38922    39658: contig of 737 in length
        39659    39758: gap of unknown length
        39759    40361: contig of 603 in length
        40362    40461: gap of unknown length
        40462    40942: contig of 481 in length
        40943    41042: gap of unknown length
        41043    41429: contig of 387 in length
        41430    41529: gap of unknown length
        41530    42050: contig of 521 in length

  Query Match             0.7%;  Score 22;  DB 2;  Length 85918;
  Best Local Similarity   100.0%;  Pred. No. 7.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    266 GGGATGCCAAAATCAATGCCCC 287
          ||||||||||||||||||||||
Db  60216 GGGATGCCAAAATCAATGCCCC 60195


RESULT 35
AC108481
LOCUS       AC108481              160986 bp    DNA     linear   INV 12-JUN-2002
DEFINITION  Drosophila melanogaster 3L BAC RP98-1019 (Roswell Park Cancer
            Institute Drosophila BAC Library) complete sequence.
ACCESSION   AC108481
VERSION     AC108481.2  GI:21392423
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 160986)
  AUTHORS   Muzny,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A.,
            Gocayne,J.D., Tabor,P., Williamson,A., Homsi,F.H.,
            Dugan-Rocha,S.D., Sodergren,E.S., Hodgson,A.H., Chen,R.C.,
            Ayele,M., Scott,G.S., Worley,K.W., Amamatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Beeson,K.Y., Brown,M., Buhay,C.,
            Busam,D.A., Center,A., Chen,G., Chen,Z., Clerc-Blankenburg,K.,
            Davenport,L.B., Dietz,S.M., Ding,Y., Dodson,K., Doup,L.E.,
            Draper,H., Emery-Cohen,A., Ferriera,S., Garg,N.D.S., Houck,J.,
            Hostin,D., Howland,T.J., Hume,J., Ibegwam,C., Jalali,M., Kovar,C.,
            Liu,W., Mattei,B., McIntosh,T.C., Morgan,M., Moy,M., Murphy,B.,
            Nelson,K.A., Ndassa,Y., Nguyen,N., Perez,L., Pittman,G.S., Puri,V.,
            Scheeler,F., Shen,H., Strong,R., Tector,C., Wang,Q., Williams,S.M.,
            Xiang,J., Zaveri,J.S., Zhou,J., Zorrilla,S., Smith,H.O.,
            Wheeler,D., Weinstock,G., Gibbs,R. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 160986)
  AUTHORS   Worley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 160986)
  AUTHORS   BCM-HGSC.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 160986)
  AUTHORS   BCM-HGSC.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 12, 2002 this sequence version replaced gi:18390133.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
```

```
            Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.
            html.
FEATURES             Location/Qualifiers
     source          1..160986
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3L"
                     /clone="RP98-1O19"
     misc_feature    56335..56355
                     /function="single clone coverage"
     misc_feature    78445
                     /note="one celera read calls extra 5 bps: CGGAC"
                     /function="polymorphic site"
     misc_feature    82770..82790
                     /note="size confimed by pcr and restriction map"
                     /function="unresolved homopolymeric run"
BASE COUNT    46643 a  34050 c  34135 g  46158 t
ORIGIN

  Query Match             0.7%;  Score 22;  DB 3;  Length 160986;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      266 GGGATGCCAAAATCAATGCCCC 287
            ||||||||||||||||||||||
Db    22969 GGGATGCCAAAATCAATGCCCC 22990


RESULT 36
AC099761
LOCUS       AC099761              267113 bp    DNA     linear   HTG 23-MAY-2002
DEFINITION  Mus musculus chromosome UNK clone RP23-230J6, WORKING DRAFT
            SEQUENCE, 65 unordered pieces.
ACCESSION   AC099761
VERSION     AC099761.2  GI:21105132
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 267113)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 267113)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2001) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 267113)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On May 23, 2002 this sequence version replaced gi:17017633.
```

```
            ---------------- Genome Center ---------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            -------------- Project Information ------------------
            Center project name: M_BA0230J06
            ----------------------------------------------------
            -------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 241118 bases at least Q40
            Consensus quality: 248108 bases at least Q30
            Consensus quality: 253552 bases at least Q20
            Insert size: 204000; agarose-fp
            Insert size: 262842; sum-of-contigs
            Quality coverage: 5.94 in Q20 bases; agarose-fp
            Quality coverage: 4.19 in Q20 bases; sum-of-contigs
            -------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 65 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1      1229: contig of 1229 bp in length
            *    1230      1329: gap of unknown length
            *    1330      2555: contig of 1226 bp in length
            *    2556      2655: gap of unknown length
            *    2656      3776: contig of 1121 bp in length
            *    3777      3876: gap of unknown length
            *    3877      5042: contig of 1166 bp in length
            *    5043      5142: gap of unknown length
            *    5143      6206: contig of 1064 bp in length
            *    6207      6306: gap of unknown length
            *    6307      7391: contig of 1085 bp in length
            *    7392      7491: gap of unknown length
            *    7492      8871: contig of 1380 bp in length
            *    8872      8971: gap of unknown length
            *    8972     10079: contig of 1108 bp in length
            *   10080     10179: gap of unknown length
            *   10180     11386: contig of 1207 bp in length
            *   11387     11486: gap of unknown length
            *   11487     12966: contig of 1480 bp in length
            *   12967     13066: gap of unknown length
            *   13067     14438: contig of 1372 bp in length
            *   14439     14538: gap of unknown length
            *   14539     16014: contig of 1476 bp in length
            *   16015     16114: gap of unknown length
            *   16115     17448: contig of 1334 bp in length
            *   17449     17548: gap of unknown length
            *   17549     18774: contig of 1226 bp in length
            *   18775     18874: gap of unknown length
            *   18875     20311: contig of 1437 bp in length
            *   20312     20411: gap of unknown length
            *   20412     21680: contig of 1269 bp in length
            *   21681     21780: gap of unknown length
            *   21781     23365: contig of 1585 bp in length
            *   23366     23465: gap of unknown length
            *   23466     24826: contig of 1361 bp in length
            *   24827     24926: gap of unknown length
            *   24927     26342: contig of 1416 bp in length
            *   26343     26442: gap of unknown length
            *   26443     27666: contig of 1224 bp in length
            *   27667     27766: gap of unknown length
            *   27767     29153: contig of 1387 bp in length
            *   29154     29253: gap of unknown length
```

```
*     29254     30291: contig of 1038 bp in length
*     30292     30391: gap of unknown length
*     30392     31735: contig of 1344 bp in length
*     31736     31835: gap of unknown length
*     31836     32976: contig of 1141 bp in length
*     32977     33076: gap of unknown length
*     33077     34547: contig of 1471 bp in length
*     34548     34647: gap of unknown length
*     34648     36327: contig of 1680 bp in length
*     36328     36427: gap of unknown length
*     36428     37876: contig of 1449 bp in length
*     37877     37976: gap of unknown length
*     37977     39593: contig of 1617 bp in length
*     39594     39693: gap of unknown length
*     39694     40839: contig of 1146 bp in length
*     40840     40939: gap of unknown length
*     40940     43016: contig of 2077 bp in length
*     43017     43116: gap of unknown length
*     43117     44627: contig of 1511 bp in length
*     44628     44727: gap of unknown length
*     44728     46220: contig of 1493 bp in length
*     46221     46320: gap of unknown length
*     46321     47684: contig of 1364 bp in length
*     47685     47784: gap of unknown length
*     47785     49420: contig of 1636 bp in length
*     49421     49520: gap of unknown length
*     49521     50957: contig of 1437 bp in length
*     50958     51057: gap of unknown length
*     51058     52514: contig of 1457 bp in length
*     52515     52614: gap of unknown length
*     52615     54905: contig of 2291 bp in length
*     54906     55005: gap of unknown length
*     55006     56149: contig of 1144 bp in length
*     56150     56249: gap of unknown length
*     56250     57770: contig of 1521 bp in length
*     57771     57870: gap of unknown length
*     57871     59283: contig of 1413 bp in length
*     59284     59383: gap of unknown length
*     59384     60656: contig of 1273 bp in length
*     60657     60756: gap of unknown length
*     60757     61903: contig of 1147 bp in length
*     61904     62003: gap of unknown length
*     62004     64818: contig of 2815 bp in length
*     64819     64918: gap of unknown length
*     64919     66167: contig of 1249 bp in length
*     66168     66267: gap of unknown length
*     66268     67997: contig of 1730 bp in length
*     67998     68097: gap of unknown length
*     68098     69365: contig of 1268 bp in length
*     69366     69465: gap of unknown length
*     69466     70770: contig of 1305 bp in length
*     70771     70870: gap of unknown length
*     70871     72937: contig of 2067 bp in length
*     72938     73037: gap of unknown length
*     73038     75049: contig of 2012 bp in length
*     75050     75149: gap of unknown length
*     75150     77650: contig of 2501 bp in length
*     77651     77750: gap of unknown length
*     77751     79482: contig of 1732 bp in length
*     79483     79582: gap of unknown length
*     79583     83087: contig of 3505 bp in length
*     83088     83187: gap of unknown length
*     83188     87699: contig of 4512 bp in length
*     87700     87799: gap of unknown length
*     87800     93702: contig of 5903 bp in length
*     93703     93802: gap of unknown length
*     93803     99216: contig of 5414 bp in length
*     99217     99316: gap of unknown length
*     99317    105763: contig of 6447 bp in length
*    105764    105863: gap of unknown length
*    105864    113706: contig of 7843 bp in length
*    113707    113806: gap of unknown length
*    113807    122387: contig of 8581 bp in length
*    122388    122487: gap of unknown length
*    122488    133894: contig of 11407 bp in length
*    133895    133994: gap of unknown length
*    133995    143826: contig of 9832 bp in length
*    143827    143926: gap of unknown length
*    143927    162163: contig of 18237 bp in length
*    162164    162263: gap of unknown length
*    162264    179812: contig of 17549 bp in length
*    179813    179912: gap of unknown length
*    179913    200350: contig of 20438 bp in length
*    200351    200450: gap of unknown length
*    200451    232908: contig of 32458 bp in length
*    232909    233008: gap of unknown length
*    233009    267113: contig of 34105 bp in length.
FEATURES             Location/Qualifiers
     source          1..267113
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-230J6"
     misc_feature    1..1229
                     /note="assembly_name:Contig35"
     misc_feature    1330..2555
                     /note="assembly_name:Contig45"
     misc_feature    2656..3776
                     /note="assembly_name:Contig55"
     misc_feature    3877..5042
                     /note="assembly_name:Contig95"
     misc_feature    5143..6206

  Query Match             0.7%;  Score 22;  DB 2;  Length 267113;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1382 TGGATGCGGGTAAGGCCATGAA 1403
           ||||||||||||||||||||||
Db  182950 TGGATGCGGGTAAGGCCATGAA 182971


RESULT 37
AE003430
LOCUS       AE003430              294218 bp    DNA     linear   INV 04-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386054 section 14
            of 35, complete sequence.
ACCESSION   AE003430 AE002566
VERSION     AE003430.2  GI:10728458
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 294218)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
```

```
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
   PUBMED   10731132
REFERENCE   2  (bases 1 to 294218)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     On Oct 9, 2000 this sequence version replaced gi:7290460.
FEATURES             Location/Qualifiers
     source          1. .294218
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="X"
     gene            complement(<3756. .>6613)
                     /gene="CG6379"
                     /map="4B1-4B1"
                     /db_xref="FLYBASE:FBan0006379"
                     /db_xref="FLYBASE:FBgn0029693"
     mRNA            complement(join(3756. .5473,5575. .>6613))
                     /gene="CG6379"
                     /product="CT19894"
                     /db_xref="FLYBASE:FBan0006379"
                     /db_xref="FLYBASE:FBgn0029693"
     CDS             complement(join(3756. .5473,5575. .6223))
                     /gene="CG6379"
                     /note="CG6379 gene product"
                     /codon_start=1
                     /protein_id="AAF45915.1"
                     /db_xref="GI:7290461"
                     /db_xref="FLYBASE:FBan0006379"
                     /db_xref="FLYBASE:FBgn0029693"
                     /translation="MDEPSDDENSEPTPKKIKREWVKSYSNKAMEMMKKMGYENDKGL
                     GKSNQGRLEPIIAVQQDGRRGFGLKLDTVQSSAGQWDPACEELEIPEPVLWLHNPGSR
                     ADAYSLDQLMGHLVTGEKKLTLDGETRYCDPAILHHILNAKTVFDDLNDNEKRRARSR
                     CNPFETIRSSIFLNRAAVKMANIDSMCNFMFTNPRDPAGQTLVAPDELLYFTDMCAGP
                     GGFSEYVLYRKSWEAKGFGFTLRGANDFKLEKFFAASPESFDTFYGVKEDGNIFDESN
                     QDSLNEYIRMHTPQGVHFAMADGGFSVEGQKNIQEILSKQLYLCQFLTALKILRPNGS
                     FVCKVFDLFTPFSVGLVYLMYKCFQQIAIIKPNSSRPANSERYLVCKYKRSDAETAGI
                     VAYLNTVNLMLSDESQLDENDVLEIFNANELAEDEDFLRYIIDSNNAIGKKQIVGLRK
                     IAAFAQNLELKETKQSEVRQECLKRWGLPDKLRQAPENKPTDRLLDELLADWANERSW
                     LSLPATEMKGVASLNATIKNVADWYFVPVGREETNINACSLFLCKSRGNLLRYTEHKK
                     WELVETAFEVQPRSIFFGQIVYEFYGEGRTIQRMAALHIIDGICLGGVDIRRRPYRER
                     VSMCDKFARSLNKPYRKDRTFGALRSKPLFRLQDMGSFFANMRHYVLKDNSQRLGFAL
                     DDNKFFVPGGIMMFCELTNNYVSAHSRSRGQLYYFNVRNKESYYKDQIPRNKADEIFA
                     SFRFSFSCRLLWKWTDLRQVEELATEDNPKILFRSDFVKFIADKLGHS"
     gene            <6990. .>8137
                     /gene="CG15375"
                     /evidence=not_experimental
                     /map="4B1-4B1"
                     /db_xref="FLYBASE:FBan0015375"
                     /db_xref="FLYBASE:FBgn0029694"
     mRNA            join(<6990. .7292,7369. .7629,7831. .7987,8046. .>8137)
                     /gene="CG15375"
                     /product="CT35417"
                     /evidence=not_experimental
                     /db_xref="FLYBASE:FBan0015375"
                     /db_xref="FLYBASE:FBgn0029694"
     CDS             join(6990. .7292,7369. .7629,7831. .7987,8046. .8137)
                     /gene="CG15375"
                     /note="CG15375 gene product"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF45916.1"
                     /db_xref="GI:7290462"
                     /db_xref="FLYBASE:FBan0015375"
                     /db_xref="FLYBASE:FBgn0029694"
                     /translation="MEEQNVDETKTNPLSRKLKFSGICKLTDEELSHLYKLDELTDEE
                     LASVDLKRSSLIDDYRLLHEVSMMRQMENEAAKPPNPPEPPKPPNPPERPKARKLLHF
                     VKVMSFEEEPGTIFTEEQLRLDSNYMELNSDIFRLESFYNYLDSMYVSGNSNRAIDLV
                     MFGIHRLNINILMNIHKFHTPKKAHQLPPRYRLIYPRGTPPTRDFTTDFSGAPGHIAM
                     LQSVLLATIFQIQLRVLAPLQPQGSQKFQAFEQQNLTVHTFRSFKIGGLKFQ"
     gene            complement(<8596. .>9779)
                     /gene="Xpac"
                     /note="CG6358"
                     /map="4B1-4B1"
                     /db_xref="FLYBASE:FBan0006358"
                     /db_xref="FLYBASE:FBgn0004832"
     mRNA            complement(join(8596. .9478,9544. .9779))
                     /gene="Xpac"
                     /product="CT19848"
                     /note="Nucleotide sequence of the Celera sequence differs
                     from the published sequence for this transcript."
                     /db_xref="FLYBASE:FBan0006358"
                     /db_xref="FLYBASE:FBgn0004832"
     CDS             complement(join(8746. .9478,9544. .9701))
                     /gene="Xpac"
                     /note="Xpac gene product; Nucleotide sequence of the
                     Celera sequence differs from the published sequence for
                     this transcript"
                     /codon_start=1
                     /protein_id="AAF45917.1"
                     /db_xref="GI:7290463"
                     /db_xref="FLYBASE:FBan0006358"
                     /db_xref="FLYBASE:FBgn0004832"
                     /translation="MSAEVSTNESAPPAEKKSKLTNAQKARIERNQAKAQKLREAKLV
                     SHPFKELASNKEGGTHPEAALSQGSSVIKVQGTKYIDSGGGFLLEQPVMPTGVGPAGL
                     NKSGEEAPPILDDAIAIPVQYEECLECGDMFADSYLFNNFGHSVCDKCRDKDERYALI
                     TRTEAKAEYLLKDCDFDKREPKLRYISRKNPHNVRWGEMKLYLHLQIHQRALEVWGSE
                     EELVRQHEAREDKREEGKARKYNKKMKQLRMEVRSSIYTKKTHEVHEHEFGPDTYDEE
                     EDTYTHTCITCPYSETYEKM"
     gene            <10068. .>11354
                     /gene="brn"
                     /note="CG4934"
                     /map="4B1-4B1"
                     /db_xref="FLYBASE:FBan0004934"
                     /db_xref="FLYBASE:FBgn0000221"
     mRNA            <10068. .11354
                     /gene="brn"
                     /product="CT15828"
                     /db_xref="FLYBASE:FBan0004934"
                     /db_xref="FLYBASE:FBgn0000221"
     CDS             10068. .11045
                     /gene="brn"
                     /note="brn gene product"
                     /codon_start=1
                     /protein_id="AAF45918.1"
                     /db_xref="GI:7290464"
                     /db_xref="FLYBASE:FBan0004934"
                     /db_xref="FLYBASE:FBgn0000221"
                     /translation="MQSKHRKLLLRCLLVLPLILLVDYCGLLTHLHELNFERHFHYPL
                     NDDTGSGSASSGLDKFAYLRVPSFTAEVPVDQPARLTMLIKSAVGNSRRREAIRRTWG
                     YEGRFSDVHLRRVFLLGTAEDSEKDVAWESREHGDILQAEFTDAYFNNTLKTMLGMRW
                     ASDQFNRSEFYLFVDDDYYVSAKNVLKFLGRGRQSHQPELLFAGHVFQTSPLRHKFSK
                     WYVSLEEYPFDRWPPYVTAGAFILSQKALRQLYAASVHLPLFRFDDVYLGIVALKAGI
                     SLQHCDDFRFHRPAYKGPDSYSSVIASHEFGDPEEMTRVWNECRSANYA"
     gene            <11717. .>16524
```

```
                     /gene="EG:EG0007.11"
                     /note="CG4944"
                     /map="4B1-4B1"
                     /db_xref="FLYBASE:FBan0004944"
                     /db_xref="FLYBASE:FBgn0026084"
     mRNA            join(11717..11855,15271..15406,15479..15592,15734..16524)
                     /gene="EG:EG0007.11"
                     /product="CT15683"
                     /db_xref="FLYBASE:FBan0004944"
                     /db_xref="FLYBASE:FBgn0026084"
     mRNA            join(12417..12509,15271..15406,15479..15592,15734..16524)
                     /gene="EG:EG0007.11"
                     /product="CT15665"
                     /db_xref="FLYBASE:FBan0004944"
                     /db_xref="FLYBASE:FBgn0026084"
     CDS             join(15271..15406,15479..15592,15734..15873)
                     /gene="EG:EG0007.11"
                     /note="EG:EG0007.11 gene product [alt 1]"
                     /codon_start=1
                     /protein_id="AAF45919.1"
                     /db_xref="GI:7290465"
                     /db_xref="FLYBASE:FBan0004944"

  Query Match             0.7%;  Score 22;  DB 3;  Length 294218;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     266 GGGATGCCAAAATCAATGCCCC 287
           ||||||||||||||||||||||
Db  158358 GGGATGCCAAAATCAATGCCCC 158379


RESULT 38
AX024141
LOCUS       AX024141                  34 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 84 from Patent FR2785293.
ACCESSION   AX024141
VERSION     AX024141.1  GI:10184452
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Nassif,X., Tinsley,C., Aujame,L., Perrin,A., Rokbi,B.,
            Bouchardon,A. and Renauld,M.G.
  JOURNAL   Patent: FR 2785293-A 84 05-MAY-2000;
            PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES             Location/Qualifiers
     source          1..34
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="amorce de"
BASE COUNT        9 a     15 c      5 g      5 t
ORIGIN

  Query Match             0.7%;  Score 21;  DB 6;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGCGAACGACCCCAACCTTC 21
           |||||||||||||||||||||
Db      14 ATGCGAACGACCCCAACCTTC 34


RESULT 39
AF281339/c
LOCUS       AF281339                 348 bp    mRNA    linear   VRT 20-FEB-2002
DEFINITION  Oncorhynchus mykiss hemopexin-like protein variant 1 mRNA, partial
            cds.
ACCESSION   AF281339
VERSION     AF281339.1  GI:11095770
KEYWORDS    .
SOURCE      Oncorhynchus mykiss.
```

```
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 348)
  AUTHORS   Bayne,C.J., Gerwick,L., Fujiki,K., Nakao,M. and Yano,T.
  TITLE     Immune-relevant (including acute phase) genes identified in the
            livers of rainbow trout, Oncorhynchus mykiss, by means of
            suppression subtractive hybridization
  JOURNAL   Dev. Comp. Immunol. 25 (3), 205-217 (2001)
  MEDLINE   21098507
   PUBMED   11164886
REFERENCE   2  (bases 1 to 348)
  AUTHORS   Bayne,C.J., Gerwick,L., Fujiki,K., Nakao,M. and Yano,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2000) Zoology, Oregon State University, 3029
            Cordley, Corvallis, OR 97331, USA
FEATURES             Location/Qualifiers
     source          1..348
                     /organism="Oncorhynchus mykiss"
                     /db_xref="taxon:8022"
     CDS             <1..>348
                     /function="binds free heme"
                     /codon_start=1
                     /product="hemopexin-like protein variant 1"
                     /protein_id="AAG30014.1"
                     /db_xref="GI:11095771"
                     /translation="NGHTFFFKGDHLWNGFKGPAQVSSAFFKELDNYHHLGHVDAAFR
                     MHNKEKPEKHDHIYFFLDDKVFSYYNHSLEEGYPKDIQLDFPGVPSHVDAAVECPKGE
                     CNSDSVLFFKGEEV"
BASE COUNT       91 a     94 c     87 g     74 t      2 others
ORIGIN

  Query Match             0.7%;  Score 21;  DB 5;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2710 CATGCGGAAGGCAGCGTCAAC 2730
           |||||||||||||||||||||
Db     135 CATGCGGAAGGCAGCGTCAAC 115


RESULT 40
HSA332365/c
LOCUS       HSA332365                768 bp    DNA     linear   PRI 18-JUL-2002
DEFINITION  Homo sapiens genomic sequence surrounding NotI site, clone
            NR5-FE23RS.
ACCESSION   AJ332365
VERSION     AJ332365.1  GI:15876783
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
            Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V.,
            Levitsky,V.G., Kolchanov,N.A., Protopopov,A.I., Kashuba,V.I.,
            Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
  TITLE     NotI flanking sequences: a tool for gene discovery and verification
            of the human genome
  JOURNAL   Nucleic Acids Res. 30 (14), 3163-3170 (2002)
   PUBMED   12136098
REFERENCE   2  (bases 1 to 768)
  AUTHORS   Zabarovsky,E.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
            Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
            Sweden
FEATURES             Location/Qualifiers
     source          1..768
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                      /clone="NR5-FE23RS"
BASE COUNT      185 a    205 c    204 g    174 t
ORIGIN

  Query Match             0.7%;  Score 21;  DB 9;  Length 768;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2420 AAACCCAACAGGACGGTGGAA 2440
          |||||||||||||||||||||
Db    537 AAACCCAACAGGACGGTGGAA 517

Search completed: January 27, 2003, 23:33:45
Job time : 10122 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 20:44:54 ; Search time 2874 Seconds
(without alignments)
18055.094 Million cell updates/sec

Title: US-09-830-433A-7
Perfect score: 3204
Sequence: 1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	25	0.8	123	17	AQ024974	AQ024974 EP(2)0842
c	2	25	0.8	365	17	BH759360	BH759360 KG02765 D
	3	25	0.8	431	13	BI241907	BI241907 RE39108.5
	4	25	0.8	440	13	BI482365	BI482365 RE65049.5
	5	25	0.8	443	9	AI063687	AI063687 GH03640.5
	6	25	0.8	485	13	BI365464	BI365464 RE50784.5
	7	25	0.8	491	13	BI165260	BI165260 RE05115.5
	8	25	0.8	521	13	BI227520	BI227520 RE24591.5
	9	25	0.8	526	13	BI363907	BI363907 RE48884.5
	10	25	0.8	531	13	BI356933	BI356933 RE43234.5
	11	25	0.8	544	9	AI114180	AI114180 GH10888.5
	12	25	0.8	545	9	AI135320	AI135320 GH13005.5
	13	25	0.8	560	13	BI362358	BI362358 RE46895.5
	14	25	0.8	577	13	BI164055	BI164055 RE03625.5
	15	25	0.8	579	13	BI165486	BI165486 RE05402.5
	16	25	0.8	583	17	AQ074074	AQ074074 EP(2)2068
c	17	25	0.8	586	17	AQ073492	AQ073492 EP(2)2323
	18	25	0.8	590	9	AI387965	AI387965 GH18695.5
	19	25	0.8	590	13	BI238617	BI238617 RE35121.5
	20	25	0.8	592	13	BI363426	BI363426 RE48231.5
	21	25	0.8	593	13	BI481998	BI481998 RE64636.5
	22	25	0.8	602	9	AI109244	AI109244 GH08375.5
	23	25	0.8	616	9	AI107819	AI107819 GH05674.5
	24	25	0.8	625	9	AI063205	AI063205 GH02745.5
	25	25	0.8	625	9	AI238498	AI238498 GH14550.5
	26	25	0.8	626	13	BI170514	BI170514 RE12003.5
	27	25	0.8	628	9	AI296017	AI296017 LP09807.5
	28	25	0.8	633	13	BI169402	BI169402 RE10566.5
	29	25	0.8	639	9	AI386994	AI386994 GH17396.5
	30	25	0.8	655	9	AI388097	AI388097 GH18879.5
	31	25	0.8	664	9	AI107422	AI107422 GH05102.5
	32	25	0.8	664	9	AI292862	AI292862 GH15875.5
	33	25	0.8	666	9	AI109291	AI109291 GH08434.5
	34	25	0.8	677	13	BI368317	BI368317 RE54268.5
	35	25	0.8	686	13	BI374631	BI374631 RE62269.5
	36	25	0.8	687	13	BI368263	BI368263 RE54204.5
	37	25	0.8	734	9	AI517363	AI517363 GH28085.5
c	38	21	0.7	496	17	AZ628307	AZ628307 1M0480O18
c	39	21	0.7	593	13	BI468065	BI468065 EST00475
	40	21	0.7	727	9	AI238153	AI238153 GH14154.5
	41	21	0.7	762	12	BF494167	BF494167 AT02596.5
	42	21	0.7	803	17	CNS05E0W	AL333113 Tetraodon
	43	21	0.7	1010	17	CNS03XQ0	AL265329 Tetraodon
c	44	21	0.7	1068	17	CNS05RXF	AL351132 Tetraodon
c	45	20	0.6	182	17	AZ577491	AZ577491 10f12 Sho
c	46	20	0.6	244	13	BI321685	BI321685 saf91d05.
c	47	20	0.6	345	12	BG588535	BG588535 EST490344
	48	20	0.6	375	17	AZ470268	AZ470268 1M0284H13
	49	20	0.6	464	14	BQ450589	BQ450589 faa79g01.
	50	20	0.6	504	9	AL817975	AL817975 AL817975
	51	20	0.6	559	17	TA27E05Q	AL453606 T. brucei
c	52	20	0.6	579	13	BM279219	BM279219 Nb_ad1_04
c	53	20	0.6	597	17	AZ214569	AZ214569 Sheared D
c	54	20	0.6	598	17	AQ941303	AQ941303 Sheared D
c	55	20	0.6	608	13	BI980081	BI980081 ft79g05.x
c	56	20	0.6	658	17	BH189926	BH189926 ATXOC42TR
	57	20	0.6	756	13	BI246500	BI246500 602958561
c	58	20	0.6	796	13	BM289442	BM289442 EST575976
	59	20	0.6	967	12	BF141559	BF141559 601788645
	60	20	0.6	1146	14	BQ734165	BQ734165 AGENCOURT
	61	19	0.6	211	12	BF758324	BF758324 CM4-CT057
	62	19	0.6	213	9	AA040282	AA040282 zk44g12.r
c	63	19	0.6	222	17	B79680	B79680 CIT-HSP-204
	64	19	0.6	255	14	R85275	R85275 yq24a05.s1
	65	19	0.6	268	10	BB563246	BB563246 BB563246
	66	19	0.6	284	9	AA172297	AA172297 zo98f04.r
	67	19	0.6	287	14	BQ571703	BQ571703 UI-M-FC0-
	68	19	0.6	305	13	BM429089	BM429089 952027F08
	69	19	0.6	313	9	AA046507	AA046507 zk67b03.r
c	70	19	0.6	315	9	AI767192	AI767192 wi94b07.x
	71	19	0.6	346	9	AI478550	AI478550 tm53d09.x
c	72	19	0.6	364	12	BG878645	BG878645 ib63g02.x
	73	19	0.6	368	10	AW016361	AW016361 UI-H-BI0-
	74	19	0.6	379	10	AW016352	AW016352 UI-H-BI0-
c	75	19	0.6	380	10	BB841859	BB841859 BB841859
	76	19	0.6	381	9	AA128881	AA128881 zm25f06.r
	77	19	0.6	396	9	AA680349	AA680349 zi11d04.s
	78	19	0.6	401	10	AW752006	AW752006 QV0-CT018
c	79	19	0.6	404	17	AQ445088	AQ445088 GSSTc0736

	80	19	0.6	412	9	AI767499	AI767499 wh27d01.x
c	81	19	0.6	418	17	AQ877258	AQ877258 HS_2142_B
	82	19	0.6	421	17	AQ088119	AQ088119 HS_2191_B
	83	19	0.6	426	14	BM712611	BM712611 UI-E-EJ0-
c	84	19	0.6	427	17	AQ966628	AQ966628 LERIL16TR
	85	19	0.6	429	10	AW469680	AW469680 hd33a04.x
	86	19	0.6	448	9	AA047592	AA047592 zf13e03.r
c	87	19	0.6	457	12	BE992116	BE992116 UI-M-BZ1-
	88	19	0.6	473	9	AI492996	AI492996 qz47e08.x
	89	19	0.6	477	10	BB852900	BB852900 BB852900
c	90	19	0.6	500	9	AL568271	AL568271 AL568271
	91	19	0.6	506	12	BG482259	BG482259 602526782
	92	19	0.6	513	17	AQ141768	AQ141768 HS_3154_B
c	93	19	0.6	517	17	AQ445307	AQ445307 GSSTc0159
c	94	19	0.6	522	17	AQ849847	AQ849847 LMAJFV1_1
c	95	19	0.6	528	12	BF540443	BF540443 602052414
c	96	19	0.6	530	17	AQ675820	AQ675820 HS_2145_A
c	97	19	0.6	533	14	BM888527	BM888527 TM198 Hum
	98	19	0.6	534	9	AA777559	AA777559 zi94d05.s
c	99	19	0.6	537	13	BG930136	BG930136 etsHEST00
	100	19	0.6	539	9	AA156683	AA156683 zl08c06.r
c	101	19	0.6	544	13	BM313985	BM313985 ih05d09.y
c	102	19	0.6	545	13	BM272641	BM272641 ih01h07.x
	103	19	0.6	550	17	AQ578653	AQ578653 nbxb0093C
	104	19	0.6	558	14	W67937	W67937 zd38g12.r1
	105	19	0.6	562	9	AI033251	AI033251 ow98e07.s
	106	19	0.6	563	17	AQ891918	AQ891918 HS_3154_B
	107	19	0.6	566	9	AI522084	AI522084 ti83e08.x
c	108	19	0.6	566	10	BE587706	BE587706 WHE0667_C
	109	19	0.6	567	13	BI771188	BI771188 603054892
c	110	19	0.6	568	13	BI726466	BI726466 1031085E0
c	111	19	0.6	570	13	BI779411	BI779411 EBro01_SQ
c	112	19	0.6	573	13	BM526937	BM526937 sa147e03.
	113	19	0.6	575	17	AZ378295	AZ378295 1M0133C06
c	114	19	0.6	578	14	BU005018	BU005018 QGG6N19.y
c	115	19	0.6	579	9	AA626793	AA626793 ad09g12.s
c	116	19	0.6	589	14	BQ776091	BQ776091 UI-H-FH0-
	117	19	0.6	592	17	BH835120	BH835120 BACPP7-F0
	118	19	0.6	596	14	BM943333	BM943333 UI-M-EH0p
c	119	19	0.6	600	13	BI991111	BI991111 4091-20 M
c	120	19	0.6	602	9	AJ447794	AJ447794 AJ447794
	121	19	0.6	608	17	BH692403	BH692403 BOHZS96TF
c	122	19	0.6	610	9	AA150480	AA150480 zl08c06.s
	123	19	0.6	611	17	AQ271191	AQ271191 nbxb0024K
	124	19	0.6	616	10	BB631390	BB631390 BB631390
	125	19	0.6	629	10	AV649619	AV649619 AV649619
c	126	19	0.6	629	10	AW375504	AW375504 QV0-CT018
	127	19	0.6	631	17	AZ995907	AZ995907 2M0281H23
c	128	19	0.6	634	9	AJ448948	AJ448948 AJ448948
c	129	19	0.6	643	9	AJ454502	AJ454502 AJ454502
c	130	19	0.6	645	12	BG414880	BG414880 HVSMEk000
c	131	19	0.6	649	9	AJ451907	AJ451907 AJ451907
	132	19	0.6	655	17	BH374446	BH374446 AG-ND-179
c	133	19	0.6	656	14	BM990290	BM990290 UI-H-DH0-
c	134	19	0.6	657	17	AQ398889	AQ398889 mgxb0011O
c	135	19	0.6	660	14	BQ573802	BQ573802 UI-H-EZ0-
	136	19	0.6	661	12	BE834042	BE834042 MR2-OT007
c	137	19	0.6	663	14	BQ573850	BQ573850 UI-H-EZ0-
c	138	19	0.6	667	14	BQ009318	BQ009318 UI-H-ED1-
c	139	19	0.6	671	14	BQ574576	BQ574576 UI-H-EZ1-
	140	19	0.6	679	14	BM714805	BM714805 UI-E-EJ0-
	141	19	0.6	682	10	AW235106	AW235106 xn18f07.x
	142	19	0.6	688	12	BG479877	BG479877 602527419
c	143	19	0.6	691	9	AJ446198	AJ446198 AJ446198
	144	19	0.6	694	12	BE731312	BE731312 601566430
c	145	19	0.6	695	14	BQ863615	BQ863615 QGC24F21.
c	146	19	0.6	698	12	BG847722	BG847722 1024019D0
c	147	19	0.6	704	14	BM977948	BM977948 UI-CF-EC1
c	148	19	0.6	710	9	AJ453198	AJ453198 AJ453198
c	149	19	0.6	714	14	BQ860993	BQ860993 QGC17D21.
c	150	19	0.6	720	14	BQ862390	BQ862390 QGC20P11.
c	151	19	0.6	725	9	AJ449948	AJ449948 AJ449948
c	152	19	0.6	730	9	AJ455348	AJ455348 AJ455348
	153	19	0.6	737	12	BF796989	BF796989 602258020
c	154	19	0.6	740	17	BH838125	BH838125 LMCR10000
c	155	19	0.6	745	14	BU005179	BU005179 QGG7F17.y
	156	19	0.6	748	13	BI087971.	BI087971 602852435
c	157	19	0.6	749	9	AJ448695	AJ448695 AJ448695
c	158	19	0.6	756	9	AJ451784	AJ451784 AJ451784
c	159	19	0.6	759	14	BQ863081	BQ863081 QGC22O18.
	160	19	0.6	767	13	BI857651	BI857651 603384966
	161	19	0.6	771	12	BE796734	BE796734 601587357
	162	19	0.6	789	10	BE544586	BE544586 601077211
	163	19	0.6	789	12	BG696610	BG696610 602658907
	164	19	0.6	793	17	BH034274	BH034274 RPCI-24-2
c	165	19	0.6	803	12	BG696444	BG696444 602659552
c	166	19	0.6	836	11	AY104690	AY104690 Zea mays
c	167	19	0.6	839	12	BF621570	BF621570 HVSMEa001
	168	19	0.6	843	17	AQ889018	AQ889018 HS_3144_A
c	169	19	0.6	850	12	BF308037	BF308037 601889419
	170	19	0.6	850	14	BQ425757	BQ425757 AGENCOURT
	171	19	0.6	852	17	CNS010ZP	AL099631 Drosophil
	172	19	0.6	854	14	BQ428170	BQ428170 AGENCOURT
	173	19	0.6	881	9	AL540182	AL540182 AL540182
	174	19	0.6	891	14	BQ423933	BQ423933 AGENCOURT
	175	19	0.6	913	9	AL557429	AL557429 AL557429
c	176	19	0.6	914	17	AZ692320	AZ692320 ENTIL19TR
	177	19	0.6	916	9	AL529762	AL529762 AL529762
	178	19	0.6	920	9	AL537237	AL537237 AL537237
c	179	19	0.6	920	9	AL563543	AL563543 AL563543
	180	19	0.6	941	12	BE783156	BE783156 601470765
	181	19	0.6	955	14	BQ918272	BQ918272 AGENCOURT
c	182	19	0.6	980	9	AL567771	AL567771 AL567771
c	183	19	0.6	988	17	CNS04BT9	AL283590 Tetraodon
	184	19	0.6	998	17	AG072512	AG072512 Pan trogl
	185	19	0.6	1006	14	BQ944799	BQ944799 AGENCOURT
	186	19	0.6	1011	9	AL542355	AL542355 AL542355
c	187	19	0.6	1036	9	AL542354	AL542354 AL542354
	188	19	0.6	1048	13	BM451556	BM451556 AGENCOURT
	189	19	0.6	1147	11	AK009713	AK009713 Mus muscu
	190	19	0.6	1166	12	BE793668	BE793668 601588747
c	191	19	0.6	1364	13	BM015155	BM015155 603641146
	192	19	0.6	1474	13	BM473239	BM473239 AGENCOURT
	193	19	0.6	1532	17	AG121472	AG121472 Pan trogl
	194	19	0.6	1924	12	BF527558	BF527558 602040481
	195	19	0.6	1942	13	BM470504	BM470504 AGENCOURT
	196	18	0.6	118	10	AW708999	AW708999 d2a11ne.r
	197	18	0.6	134	14	F25511	F25511 HSPD12591 H
c	198	18	0.6	167	10	BB563762	BB563762 BB563762
	199	18	0.6	171	13	BI764138	BI764138 603043368
c	200	18	0.6	171	17	AZ285464	AZ285464 RPCI-23-4
	201	18	0.6	172	13	BI771966	BI771966 603058996
	202	18	0.6	181	13	BI824405	BI824405 603038637
	203	18	0.6	185	13	BI915403	BI915403 603177519
	204	18	0.6	190	10	AW721882	AW721882 a1c12nm.r
	205	18	0.6	192	10	AW711311	AW711311 f1d07ne.r
	206	18	0.6	207	9	AI319959	AI319959 b7f02nm.r
	207	18	0.6	211	10	AW879740	AW879740 QV3-OT002
	208	18	0.6	223	10	BB591304	BB591304 BB591304
	209	18	0.6	227	9	AV253227	AV253227 AV253227
	210	18	0.6	230	10	AW254220	AW254220 UI-R-BJ0-
c	211	18	0.6	241	10	AW222147	AW222147 EST298958
	212	18	0.6	246	9	AI319735	AI319735 b5c09nm.r
	213	18	0.6	247	12	BG209826	BG209826 RST29354
	214	18	0.6	247	13	BG933351	BG933351 WS1_3_E10
c	215	18	0.6	253	17	AZ561628	AZ561628 RPCI-23-1
	216	18	0.6	254	10	AW722970	AW722970 c9h12nm.r
c	217	18	0.6	255	12	BG788767	BG788767 SEAUMC008
c	218	18	0.6	256	9	AI426086	AI426086 mh36a03.x
c	219	18	0.6	258	9	AA715375	AA715375 nv52a08.r
c	220	18	0.6	258	12	BF290318	BF290318 EST454909
	221	18	0.6	263	10	AW709370	AW709370 d4h03ne.r
	222	18	0.6	265	9	AL840777	AL840777 AL840777
c	223	18	0.6	268	10	BE211550	BE211550 so63c01.y
	224	18	0.6	271	10	AV629344	AV629344 AV629344
c	225	18	0.6	273	10	BE425248	BE425248 WHE313_A0

c	226	18	0.6	279	10	BB564822	BB564822 BB564822
	227	18	0.6	280	9	AI176829	AI176829 EST220429
c	228	18	0.6	283	10	AW229465	AW229465 uo59c06.y
	229	18	0.6	284	10	AW718106	AW718106 i3g02nm.r
	230	18	0.6	284	10	BB045431	BB045431 BB045431
c	231	18	0.6	286	17	CNS03TK6	AL259935 Tetraodon
	232	18	0.6	292	10	AW708555	AW708555 c7d11ne.f
	233	18	0.6	294	9	AI320261	AI320261 c2e12nm.r
	234	18	0.6	297	13	BI999212	BI999212 1031071G0
c	235	18	0.6	298	10	BE341778	BE341778 EST394610
c	236	18	0.6	300	14	R53424	R53424 yj70h06.r1
c	237	18	0.6	303	12	BF754561	BF754561 QV3-CT056
c	238	18	0.6	305	12	BE950617	BE950617 UI-M-CE0-
c	239	18	0.6	309	9	AA031134	AA031134 mi47b09.r
	240	18	0.6	312	13	BI313612	BI313612 dad80f10.
c	241	18	0.6	313	9	AI581489	AI581489 ar63d10.x
c	242	18	0.6	314	9	AI639612	AI639612 vv50c02.x
	243	18	0.6	314	10	AW718452	AW718452 i5d09nm.r
	244	18	0.6	315	9	AU065636	AU065636 AU065636
	245	18	0.6	316	14	BM793718	BM793718 K-EST0074
	246	18	0.6	317	17	AZ004714	AZ004714 RPCI-23-3
c	247	18	0.6	319	10	BE646187	BE646187 7e82g03.x
	248	18	0.6	321	10	AW796612	AW796612 CM1-UM003
c	249	18	0.6	326	10	AW139530	AW139530 UI-H-BI1-
	250	18	0.6	326	10	AW711619	AW711619 f4e08ne.r
	251	18	0.6	327	17	BH233280	BH233280 1006172G0
c	252	18	0.6	328	13	BG950118	BG950118 CM2-CT063
	253	18	0.6	335	13	BI305770	BI305770 NL_1_K18
c	254	18	0.6	340	9	AA936118	AA936118 om57a12.s
	255	18	0.6	340	14	F18818	F18818 HSPD03415 H
	256	18	0.6	341	10	AW858006	AW858006 PM2-CT032
c	257	18	0.6	343	9	AA908395	AA908395 og77d07.s
c	258	18	0.6	343	9	AI895615	AI895615 EST265058
c	259	18	0.6	345	14	BM755626	BM755626 K-EST0033
c	260	18	0.6	346	12	BF731441	BF731441 mab86e10.
	261	18	0.6	346	17	AQ105704	AQ105704 HS_3048_B
	262	18	0.6	347	10	BE475731	BE475731 946048B02
	263	18	0.6	347	12	BF290317	BF290317 EST454908
c	264	18	0.6	347	13	BI134217	BI134217 UI-M-BH3-
c	265	18	0.6	351	12	BF908639	BF908639 IL2-UT007
c	266	18	0.6	352	9	AI918044	AI918044 tt84h03.x
	267	18	0.6	352	12	BF155936	BF155936 CM2-HT094
	268	18	0.6	352	17	BH233263	BH233263 1006172F1
	269	18	0.6	356	12	BF954591	BF954591 MR4-NN020
	270	18	0.6	359	9	AA289795	AA289795 mp12d07.r
c	271	18	0.6	359	10	BB840724	BB840724 BB840724
c	272	18	0.6	359	13	BI550530	BI550530 603195454
	273	18	0.6	361	12	BF401635	BF401635 UI-R-CA0-
c	274	18	0.6	363	12	BF909518	BF909518 IL2-UT007
	275	18	0.6	363	17	AZ785204	AZ785204 2M0028O21
	276	18	0.6	366	10	AW715053	AW715053 g1g10nm.r
	277	18	0.6	366	10	AW722189	AW722189 a5g05nm.r
	278	18	0.6	366	10	AW722424	AW722424 b3a06nm.r
	279	18	0.6	366	10	AW722437	AW722437 b3b08nm.r
	280	18	0.6	366	10	AW924967	AW924967 WS1_73_G0
	281	18	0.6	366	17	BH233207	BH233207 1006172D0
c	282	18	0.6	367	9	AU229563	AU229563 AU229563
	283	18	0.6	368	10	AW722960	AW722960 c9g08nm.r
c	284	18	0.6	370	9	AI899318	AI899318 EST268761
	285	18	0.6	375	9	AI320908	AI320908 d2a02nm.r
c	286	18	0.6	375	12	BE717619	BE717619 RC4-HT078
	287	18	0.6	376	12	BF880542	BF880542 QV3-ET019
	288	18	0.6	377	14	BM770306	BM770306 K-EST0053
	289	18	0.6	378	14	T55187	T55187 yb43e06.s1
	290	18	0.6	379	13	BM146604	BM146604 TCAAP1E64
	291	18	0.6	379	17	BH228099	BH228099 1006144C0
c	292	18	0.6	381	10	BE645301	BE645301 7e67h02.x
c	293	18	0.6	382	10	AW244249	AW244249 BR_END01D
	294	18	0.6	383	9	AI673552	AI673552 we75d09.x
c	295	18	0.6	384	12	BG556855	BG556855 EM1_39_E0
c	296	18	0.6	386	9	AI140320	AI140320 qa36h05.x
	297	18	0.6	387	10	AW715932	AW715932 g8a12nm.r
	298	18	0.6	387	12	BF293691	BF293691 WHE2156_h
	299	18	0.6	390	10	AW715157	AW715157 g2e12nm.r
	300	18	0.6	390	10	AW718187	AW718187 i4b12nm.r
	301	18	0.6	390	10	AW722119	AW722119 a4e08nm.r
c	302	18	0.6	392	9	AI274846	AI274846 ql73c06.x
	303	18	0.6	392	14	D49236	D49236 RICS16373A
c	304	18	0.6	394	9	AA992621	AA992621 ot81a01.s
	305	18	0.6	394	17	AQ489855	AQ489855 RPCI-11-2
c	306	18	0.6	395	9	AI646144	AI646144 uc41e03.x
c	307	18	0.6	395	14	H46968	H46968 yo19g09.r1
	308	18	0.6	399	9	AI327735	AI327735 i0c06a1.f
c	309	18	0.6	399	12	BG135223	BG135223 EST468115
c	310	18	0.6	400	9	AA618058	AA618058 nq02g12.s
c	311	18	0.6	401	9	AA488254	AA488254 ad08f02.r
	312	18	0.6	401	14	BM824436	BM824436 K-EST0095
	313	18	0.6	402	9	AA848564	AA848564 EST191324
c	314	18	0.6	402	9	AA254573	AA254573 va13d07.r
	315	18	0.6	404	9	AA303191	AA303191 EST15945
c	316	18	0.6	404	12	BG097096	BG097096 EST461615
c	317	18	0.6	405	9	AA663519	AA663519 ab80c09.s
c	318	18	0.6	406	9	AI265847	AI265847 qq11e01.x
c	319	18	0.6	409	10	AW406347	AW406347 UI-HF-BL0
c	320	18	0.6	409	12	BG736828	BG736828 rk63a05.y
c	321	18	0.6	411	9	AI352374	AI352374 qt21a04.x
c	322	18	0.6	411	17	AQ420481	AQ420481 RPCI-11-1
c	323	18	0.6	413	10	BB688204	BB688204 BB688204
	324	18	0.6	414	9	AI700328	AI700328 wd06d04.x
c	325	18	0.6	414	10	AV640008	AV640008 AV640008
	326	18	0.6	415	17	B91646	B91646 CIT-HSP-217
c	327	18	0.6	416	9	AA722666	AA722666 zg90g07.s
	328	18	0.6	416	9	AI508201	AI508201 mh36a03.y
c	329	18	0.6	417	14	W31201	W31201 zb64h08.r1
c	330	18	0.6	419	14	H40724	H40724 yo08b08.r1
	331	18	0.6	420	10	AW009754	AW009754 ws87a05.x
c	332	18	0.6	421	9	AA455452	AA455452 aa17g03.s
c	333	18	0.6	421	10	AW296946	AW296946 UI-H-BW0-
c	334	18	0.6	421	10	BB863755	BB863755 BB863755
c	335	18	0.6	422	9	AA226545	AA226545 nc26b11.r
c	336	18	0.6	424	9	AI827283	AI827283 wk90g03.x
c	337	18	0.6	424	10	BB864748	BB864748 BB864748
	338	18	0.6	426	10	AW722015	AW722015 a3b04nm.r
	339	18	0.6	426	17	AZ154921	AZ154921 SP_0032_A
c	340	18	0.6	427	9	AA542249	AA542249 vk28c08.r
c	341	18	0.6	427	14	W74161	W74161 zd75h08.s1
c	342	18	0.6	427	17	AQ023109	AQ023109 HS_2182_B
c	343	18	0.6	428	14	W31772	W31772 zb94h03.r1
	344	18	0.6	429	10	AV657597	AV657597 AV657597
c	345	18	0.6	430	9	AA512922	AA512922 nh91a03.s
c	346	18	0.6	430	12	BG736762	BG736762 rk62a09.y
	347	18	0.6	431	10	AW711485	AW711485 f3a04ne.r
c	348	18	0.6	431	17	AQ400292	AQ400292 072_H_12-
c	349	18	0.6	432	12	BG456837	BG456837 NF098F05P
c	350	18	0.6	432	17	AZ425529	AZ425529 1M0205J08
c	351	18	0.6	433	9	AI074905	AI074905 oy04b01.s
c	352	18	0.6	434	14	R05906	R05906 ye93g08.r1
	353	18	0.6	435	9	AA016351	AA016351 mh36a03.r
c	354	18	0.6	435	9	AA078878	AA078878 zm21f11.s
c	355	18	0.6	436	9	AI735032	AI735032 as68h01.x
c	356	18	0.6	437	10	AW592483	AW592483 hf43c01.x
	357	18	0.6	438	9	AA226209	AA226209 nc26b11.s
	358	18	0.6	438	14	BM824375	BM824375 K-EST0095
c	359	18	0.6	438	17	BH181601	BH181601 019_K_05-
c	360	18	0.6	438	17	CNS07N6T	AL618551 T7 end of
c	361	18	0.6	439	9	AA787018	AA787018 k5h03a1.f
	362	18	0.6	439	9	AA455451	AA455451 aa17g03.r
c	363	18	0.6	440	13	BJ260035	BJ260035 BJ260035
	364	18	0.6	441	13	BM659375	BM659375 SGD602702
c	365	18	0.6	442	9	AA740463	AA740463 ob24c09.s
c	366	18	0.6	444	14	BQ596235	BQ596235 PfESToab3
c	367	18	0.6	445	9	AA619148	AA619148 vo68f02.r
c	368	18	0.6	445	10	BB862118	BB862118 BB862118
c	369	18	0.6	446	9	AI892708	AI892708 ms02h02.y
c	370	18	0.6	446	13	BI290088	BI290088 UI-R-DK0-
c	371	18	0.6	447	12	BG604563	BG604563 WHE0947_C

c	372	18	0.6	449	9	AI989401	AI989401 wt80g06.x
	373	18	0.6	450	9	AI744144	AI744144 wc36g09.x
c	374	18	0.6	450	9	AA596973	AA596973 vo16g07.r
	375	18	0.6	450	17	BH233160	BH233160 1006172A0
c	376	18	0.6	451	9	AI317044	AI317044 uk71g05.y
c	377	18	0.6	451	9	AI810755	AI810755 tu04c02.x
c	378	18	0.6	452	9	AI358863	AI358863 qy24a04.x
c	379	18	0.6	453	10	AV866472	AV866472 AV866472
c	380	18	0.6	455	9	AA848571	AA848571 EST191331
	381	18	0.6	455	9	AU174161	AU174161 AU174161
c	382	18	0.6	456	10	BB859431	BB859431 BB859431
	383	18	0.6	457	17	AQ936179	AQ936179 RPCI-11-S
c	384	18	0.6	458	10	AW049881	AW049881 UI-M-BH1-
	385	18	0.6	458	10	BB751397	BB751397 BB751397
c	386	18	0.6	458	10	BB826918	BB826918 BB826918
c	387	18	0.6	458	10	BE630893	BE630893 uu50e02.x
c	388	18	0.6	459	17	BH624494	BH624494 1007078G0
c	389	18	0.6	460	9	AA309020	AA309020 EST179803
	390	18	0.6	461	10	AW279182	AW279182 sf67e08.y
c	391	18	0.6	461	14	BM852619	BM852619 K-EST0133
c	392	18	0.6	462	9	AL527075	AL527075 AL527075
	393	18	0.6	462	10	BE018547	BE018547 bb81h06.y
c	394	18	0.6	462	10	BE463060	BE463060 EST354355
	395	18	0.6	463	9	AU091433	AU091433 AU091433
c	396	18	0.6	463	13	BI098513	BI098513 IP1_33_D0
c	397	18	0.6	464	12	BF011937	BF011937 us37h06.y
c	398	18	0.6	464	12	BF117399	BF117399 uz36f01.y
c	399	18	0.6	465	10	BE471072	BE471072 WHE0283_C
c	400	18	0.6	466	9	AA488220	AA488220 ad08b02.r
c	401	18	0.6	467	10	AW222105	AW222105 EST298916
c	402	18	0.6	468	14	N42261	N42261 yw85e05.r1
c	403	18	0.6	468	17	AZ424608	AZ424608 1M0204J15
c	404	18	0.6	470	9	AA797071	AA797071 vs95h12.r
	405	18	0.6	471	12	BG317135	BG317135 947025G04
c	406	18	0.6	475	12	BF564561	BF564561 UI-R-C4-a
c	407	18	0.6	475	12	BG143201	BG143201 ia94h08.y
c	408	18	0.6	475	13	BI052727	BI052727 RC0-GN027
	409	18	0.6	476	10	AV626273	AV626273 AV626273
	410	18	0.6	481	9	AA111795	AA111795 SW3ICA189
	411	18	0.6	483	9	AI340161	AI340161 qo79b05.x
	412	18	0.6	483	9	AA526293	AA526293 ni95d06.s
c	413	18	0.6	483	9	AA576035	AA576035 nm57f07.s
	414	18	0.6	483	14	BM700236	BM700236 UI-E-DW1-
	415	18	0.6	484	9	AA745478	AA745478 ny64a05.s
	416	18	0.6	484	9	AA910358	AA910358 ok83f12.s
	417	18	0.6	484	10	AW722681	AW722681 c4h12nm.r
c	418	18	0.6	484	13	BM611036	BM611036 170006591
	419	18	0.6	485	9	AA504330	AA504330 aa61e10.s
	420	18	0.6	487	17	BH217098	BH217098 1006051F0
c	421	18	0.6	489	9	AI188204	AI188204 qd66f12.x
c	422	18	0.6	489	14	W34918	W34918 mc60d10.r1
	423	18	0.6	490	12	BF436933	BF436933 7p66c02.x
c	424	18	0.6	490	17	BH627981	BH627981 1007078G0
c	425	18	0.6	491	10	BE334969	BE334969 us91c04.y
	426	18	0.6	492	10	BE518399	BE518399 WHE0815_F
	427	18	0.6	495	9	AI636631	AI636631 ts92a04.x
c	428	18	0.6	495	9	AI898139	AI898139 EST267582
c	429	18	0.6	495	9	AI945998	AI945998 bs19d03.y
	430	18	0.6	495	10	AW535515	AW535515 UI-R-C4-a
c	431	18	0.6	495	13	BI725114	BI725114 1031077A1
c	432	18	0.6	496	9	AA823043	AA823043 vw39h11.r
c	433	18	0.6	496	14	BQ168065	BQ168065 WHE0031.E
c	434	18	0.6	497	10	BB713501	BB713501 BB713501
	435	18	0.6	499	9	AI529922	AI529922 ui87a12.y
	436	18	0.6	499	10	AW722665	AW722665 c4f10nm.r
	437	18	0.6	499	12	BG396040	BG396040 602458779
c	438	18	0.6	499	17	AZ033668	AZ033668 RPCI-23-3
c	439	18	0.6	501	9	AI897634	AI897634 EST267077
	440	18	0.6	502	10	BE481701	BE481701 167173 BA
c	441	18	0.6	504	10	BE587912	BE587912 WHE0665_C
	442	18	0.6	508	13	BI718806	BI718806 1031031F0
c	443	18	0.6	508	14	BM799338	BM799338 K-EST0083
c	444	18	0.6	510	10	AW762147	AW762147 ur54c05.y
	445	18	0.6	510	13	BM123923	BM123923 L0532G03-
	446	18	0.6	512	10	BE601703	BE601703 HVSMEh009
	447	18	0.6	515	17	AQ173539	AQ173539 HS_3193_A
	448	18	0.6	517	12	BG383572	BG383572 301725 MA
c	449	18	0.6	517	12	BF257675	BF257675 HVSMEf001
	450	18	0.6	517	14	BQ241055	BQ241055 TaE05009F
	451	18	0.6	519	10	BE453917	BE453917 946048B02
c	452	18	0.6	520	12	BG131438	BG131438 EST464330
c	453	18	0.6	521	9	AA003883	AA003883 mg79c08.r
c	454	18	0.6	521	9	AI684218	AI684218 tx80a05.x
	455	18	0.6	523	17	AQ592249	AQ592249 HS_5331_B
	456	18	0.6	524	10	BE415327	BE415327 MWL027.E0
	457	18	0.6	525	14	BM792906	BM792906 K-EST0073
	458	18	0.6	528	17	BH217120	BH217120 1006051H0
c	459	18	0.6	529	10	AW341623	AW341623 hd12c06.x
c	460	18	0.6	529	10	BE445703	BE445703 WHE1452_H
	461	18	0.6	529	12	BG312601	BG312601 WHE2455_F
c	462	18	0.6	531	12	BG135151	BG135151 EST468043
c	463	18	0.6	532	10	BE444831	BE444831 WHE1129_C
c	464	18	0.6	532	10	BE471739	BE471739 EST416592
c	465	18	0.6	533	12	BF098566	BF098566 EST429087
c	466	18	0.6	533	17	AQ468424	AQ468424 HS_5211_B
	467	18	0.6	534	12	BG316633	BG316633 947021F08
	468	18	0.6	535	10	BE220391	BE220391 hv71f08.x
c	469	18	0.6	535	12	BG396112	BG396112 602458867
c	470	18	0.6	535	13	BJ208920	BJ208920 BJ208920
	471	18	0.6	535	14	BQ810173	BQ810173 1030015H0
c	472	18	0.6	536	10	BE443512	BE443512 WHE1115_D
c	473	18	0.6	537	9	AA609344	AA609344 zu71b09.s
c	474	18	0.6	537	9	AA609958	AA609958 af09f03.s
	475	18	0.6	537	12	BE750691	BE750691 202174 MA
	476	18	0.6	537	14	BQ186092	BQ186092 UI-E-EJ1-
	477	18	0.6	538	12	BF266267	BF266267 HV_CEa001
	478	18	0.6	540	12	BG410486	BG410486 947048E10
	479	18	0.6	542	10	AW861661	AW861661 PM2-CT032
c	480	18	0.6	542	12	BG591622	BG591622 EST499464
	481	18	0.6	542	17	BH233474	BH233474 1006173H1
	482	18	0.6	543	9	AI189604	AI189604 qd32f08.x
	483	18	0.6	543	10	BE494118	BE494118 WHE1251_H
c	484	18	0.6	544	13	BI747628	BI747628 rm42h07.y
	485	18	0.6	545	10	AW857994	AW857994 PM2-CT032
	486	18	0.6	545	13	BI542714	BI542714 949021G05
	487	18	0.6	546	9	AA795881	AA795881 vt28g03.r
	488	18	0.6	546	9	AI279507	AI279507 ql51b10.x
c	489	18	0.6	546	9	AI296467	AI296467 LP10381.5
c	490	18	0.6	546	12	BF607495	BF607495 MY1_00043
	491	18	0.6	546	12	BG312656	BG312656 WHE2456_D
c	492	18	0.6	547	10	AW623020	AW623020 EST320965
c	493	18	0.6	547	10	BE586853	BE586853 WHE0508_C
	494	18	0.6	547	12	BF073278	BF073278 219888 MA
c	495	18	0.6	548	12	BG887745	BG887745 EST513596
c	496	18	0.6	549	9	AI608132	AI608132 va13d07.y
	497	18	0.6	551	17	AZ344175	AZ344175 1M0078J03
c	498	18	0.6	552	12	BF586260	BF586260 FM1_27_F1
	499	18	0.6	554	12	BF320458	BF320458 uz53e12.y
	500	18	0.6	557	17	TA277H10Q	AL485754 T. brucei
	501	18	0.6	558	10	AW747579	AW747579 WS1_73_G0
	502	18	0.6	558	12	BG191235	BG191235 RST10322
c	503	18	0.6	558	13	BM664995	BM664995 UI-E-CQ1-
	504	18	0.6	558	17	BH116351	BH116351 RPCI-24-2
c	505	18	0.6	560	13	BI181668	BI181668 UNL-P-FN-
c	506	18	0.6	560	13	BI184776	BI184776 UNL-P-FN-
	507	18	0.6	560	14	N36019	N36019 yy01a10.r1
	508	18	0.6	561	9	AI186453	AI186453 qd26b12.x
	509	18	0.6	561	17	BH769593	BH769593 BMBAC366F
c	510	18	0.6	562	10	BE443310	BE443310 WHE1112_F
	511	18	0.6	563	9	AJ443060	AJ443060 AJ443060
	512	18	0.6	564	9	AI194352	AI194352 ui14f09.r
	513	18	0.6	564	17	AQ837843	AQ837843 HS_4647_B
c	514	18	0.6	566	13	BG947217	BG947217 IP1_3_H10
c	515	18	0.6	568	13	BI996621	BI996621 1031040F0
c	516	18	0.6	569	13	BM191050	BM191050 daj31d03.
	517	18	0.6	569	17	AZ617784	AZ617784 1M0449P11

518	18	0.6	570	12	BG815845	BG815845 dad74c11.
c 519	18	0.6	572	9	AA031174	AA031174 mi48f09.r
c 520	18	0.6	572	14	W83563	W83563 mf29c10.r1
521	18	0.6	573	10	BE352019	BE352019 894054G03
522	18	0.6	573	12	BF732233	BF732233 nae08a05.
523	18	0.6	575	10	BE426467	BE426467 WHE0335_D
524	18	0.6	575	12	BG446435	BG446435 GA__Eb003
c 525	18	0.6	575	14	C78684	C78684 C78684 Mous
c 526	18	0.6	576	9	AA021851	AA021851 mh82c10.r
527	18	0.6	576	12	BG285981	BG285981 602381158
528	18	0.6	579	13	BI721597	BI721597 1031056H0
c 529	18	0.6	579	17	AQ449450	AQ449450 mgxb0023J
530	18	0.6	581	12	BF429279	BF429279 WHE1801_F
531	18	0.6	581	14	W30302	W30302 mc28d03.r1
c 532	18	0.6	582	9	AA510905	AA510905 vh60b11.r
c 533	18	0.6	583	9	AI833790	AI833790 605095D11
534	18	0.6	583	14	BM839874	BM839874 K-EST0116
535	18	0.6	584	12	BF035496	BF035496 601454788
536	18	0.6	586	13	BI360436	BI360436 387343 MA
537	18	0.6	588	12	BG086714	BG086714 H3129F11-
c 538	18	0.6	588	14	BQ117267	BQ117267 EST602843
539	18	0.6	588	14	BQ554291	BQ554291 H4027C04-
c 540	18	0.6	589	12	BG875929	BG875929 IL3-CT022
541	18	0.6	590	9	AA009204	AA009204 mh02e02.r
c 542	18	0.6	590	9	AA273515	AA273515 vc38a02.r
c 543	18	0.6	590	12	BE759313	BE759313 an_3667 A
c 544	18	0.6	591	12	BG821417	BG821417 602724873
c 545	18	0.6	592	9	AU144868	AU144868 AU144868
546	18	0.6	592	12	BF037105	BF037105 601456871
c 547	18	0.6	592	12	BF177019	BF177019 EM1_2_E09
c 548	18	0.6	595	9	AI296538	AI296538 LP10481.5
c 549	18	0.6	596	14	C78675	C78675 C78675 Mous
c 550	18	0.6	597	10	AW211329	AW211329 uo79f10.y
c 551	18	0.6	597	12	BG333464	BG333464 602460113
552	18	0.6	598	17	AZ970992	AZ970992 2M0244O13
c 553	18	0.6	600	9	AA275796	AA275796 vc25h11.r
c 554	18	0.6	600	12	BG801359	BG801359 0102-57 M
c 555	18	0.6	600	13	BI991681	BI991681 4104-04 M
c 556	18	0.6	601	13	BI728601	BI728601 1031100C0
557	18	0.6	605	12	BF702701	BF702701 MI-P-E3-a
c 558	18	0.6	606	10	AW961258	AW961258 EST373330
c 559	18	0.6	606	14	BQ564654	BQ564654 gi20f12.y
c 560	18	0.6	608	9	AA239467	AA239467 my20a08.r
c 561	18	0.6	609	13	BI144499	BI144499 602908688
c 562	18	0.6	609	17	AZ269297	AZ269297 RPCI-23-7
563	18	0.6	609	17	AG161306	AG161306 Pan trogl
c 564	18	0.6	611	13	BM670515	BM670515 UI-E-DW1-
c 565	18	0.6	612	10	AW741075	AW741075 uq94e06.y
566	18	0.6	615	9	AL680055	AL680055 AL680055
c 567	18	0.6	617	9	AI519863	AI519863 LD39687.5
c 568	18	0.6	617	10	BB616288	BB616288 BB616288
569	18	0.6	619	9	AL628932	AL628932 AL628932
570	18	0.6	620	10	BE392600	BE392600 601307424
c 571	18	0.6	621	12	BF153194	BF153194 010F07 Ma
c 572	18	0.6	621	13	BJ096108	BJ096108 BJ096108
c 573	18	0.6	622	12	BG130752	BG130752 EST463644
574	18	0.6	623	10	BE390292	BE390292 601285181
c 575	18	0.6	623	10	BE435420	BE435420 EST406498
c 576	18	0.6	623	10	BE443048	BE443048 WHE1114_B
c 577	18	0.6	624	13	BI110309	BI110309 602900740
c 578	18	0.6	626	9	AI667826	AI667826 605028H07
c 579	18	0.6	626	9	AI947045	AI947045 bs34d08.y
580	18	0.6	626	17	AZ000930	AZ000930 RPCI-23-3
581	18	0.6	627	10	BB656104	BB656104 BB656104
c 582	18	0.6	628	9	AI297117	AI297117 LP11337.5
c 583	18	0.6	628	9	AU215688	AU215688 AU215688
584	18	0.6	628	10	BE494858	BE494858 WHE1258_E
c 585	18	0.6	628	13	BG915757	BG915757 602814288
c 586	18	0.6	629	10	AW414162	AW414162 uo93f01.y
c 587	18	0.6	629	13	BM411907	BM411907 EST586234
588	18	0.6	632	10	BB662463	BB662463 BB662463
589	18	0.6	632	12	BG031712	BG031712 602300868
c 590	18	0.6	632	13	BI725113	BI725113 1031077A1
591	18	0.6	632	13	BI855623	BI855623 603383128
592	18	0.6	632	14	BM949826	BM949826 UI-M-EG0p
c 593	18	0.6	634	13	BG922229	BG922229 602820012
c 594	18	0.6	634	13	BI221817	BI221817 602937080
c 595	18	0.6	635	10	AW565147	AW565147 LG1_328_H
596	18	0.6	636	9	AI734328	AI734328 606029F11
597	18	0.6	636	9	AL644649	AL644649 AL644649
c 598	18	0.6	636	13	BI718554	BI718554 1031028E0
c 599	18	0.6	638	12	BF164209	BF164209 601773155
600	18	0.6	639	10	BB650274	BB650274 BB650274
c 601	18	0.6	639	10	BB661282	BB661282 BB661282
c 602	18	0.6	642	13	BM642443	BM642443 170006873
603	18	0.6	643	10	BE407535	BE407535 601300382
c 604	18	0.6	645	17	AZ337868	AZ337868 1M0068O11
605	18	0.6	645	17	BH399386	BH399386 AG-ND-140
c 606	18	0.6	647	9	AU236697	AU236697 AU236697
c 607	18	0.6	647	13	BI919702	BI919702 EST539637
608	18	0.6	647	17	AQ938910	AQ938910 NL1-DL8R
609	18	0.6	650	17	B26224	B26224 T5N8TFB TAM
610	18	0.6	651	12	BG723040	BG723040 602695579
611	18	0.6	652	12	BF116020	BF116020 7n76d10.x
612	18	0.6	653	17	BH030656	BH030656 RPCI-24-2
c 613	18	0.6	654	10	BE389861	BE389861 601284368
c 614	18	0.6	655	17	AZ371972	AZ371972 1M0123E11
c 615	18	0.6	656	10	BB291771	BB291771 BB291771
c 616	18	0.6	656	12	BG483701	BG483701 602503556
c 617	18	0.6	659	10	AW221934	AW221934 EST298745
618	18	0.6	660	17	CNS02557	AL181636 Tetraodon
c 619	18	0.6	661	13	BG966873	BG966873 602834374
620	18	0.6	662	13	BI193833	BI193833 602948330
c 621	18	0.6	662	13	BM592388	BM592388 170006874
c 622	18	0.6	663	12	BG453166	BG453166 NF089B04L
c 623	18	0.6	664	14	BQ258352	BQ258352 NISC_kp12
c 624	18	0.6	665	14	BM938713	BM938713 UI-M-CG0p
625	18	0.6	665	17	AZ392170	AZ392170 1M0154K16
c 626	18	0.6	667	10	BE602091	BE602091 HVSMEh010
627	18	0.6	668	12	BG703583	BG703583 602686419
c 628	18	0.6	668	13	BJ497081	BJ497081 BJ497081
c 629	18	0.6	669	12	BG598954	BG598954 EST503854
c 630	18	0.6	670	14	BM977765	BM977765 UI-CF-EN1
631	18	0.6	670	17	BH004891	BH004891 BMBAC08M1
c 632	18	0.6	670	17	AG127148	AG127148 Pan trogl
c 633	18	0.6	671	10	BB414804	BB414804 BB414804
c 634	18	0.6	672	14	BQ019716	BQ019716 UI-H-ED0-
635	18	0.6	674	10	BB554612	BB554612 BB554612
c 636	18	0.6	675	14	BQ806921	BQ806921 WHE3584_F
637	18	0.6	677	14	BQ411837	BQ411837 GA__Ed004
c 638	18	0.6	677	14	BQ473104	BQ473104 rc84a11.y
c 639	18	0.6	677	14	BQ837962	BQ837962 WHE2905_A
640	18	0.6	679	17	AZ095325	AZ095325 RPCI-23-1
c 641	18	0.6	680	17	AZ573715	AZ573715 320PvE06
c 642	18	0.6	682	13	BM321369	BM321369 rockefell
643	18	0.6	683	9	AL662455	AL662455 AL662455
644	18	0.6	683	13	BG923134	BG923134 602823962
c 645	18	0.6	683	13	BJ136784	BJ136784 BJ136784
c 646	18	0.6	683	17	AZ323647	AZ323647 1M0045M10
c 647	18	0.6	685	9	AU213990	AU213990 AU213990
648	18	0.6	685	10	BB042440	BB042440 BB042440
649	18	0.6	685	12	BF117815	BF117815 uz08c08.y
650	18	0.6	685	14	BU005974	BU005974 QGG9J14.y
c 651	18	0.6	686	14	BQ789194	BQ789194 WHE4158_F
652	18	0.6	687	13	BM448830	BM448830 DSA028G02
c 653	18	0.6	688	14	W76810	W76810 me83f07.r1
c 654	18	0.6	692	13	BI552119	BI552119 603195061
655	18	0.6	693	10	BE378712	BE378712 601237147
656	18	0.6	694	14	BM840013	BM840013 K-EST0117
657	18	0.6	697	14	BM680291	BM680291 UI-E-EO1-
658	18	0.6	701	10	AV751438	AV751438 AV751438
c 659	18	0.6	701	13	BI151263	BI151263 602917231
c 660	18	0.6	701	17	AZ951938	AZ951938 2M0216N04
c 661	18	0.6	702	12	BG764245	BG764245 602736116
662	18	0.6	702	17	BH385679	BH385679 AG-ND-140
c 663	18	0.6	705	13	BI560864	BI560864 603254033

	664	18	0.6	705	13	BM020234	BM020234 603648828
c	665	18	0.6	706	13	BI410583	BI410583 602964143
c	666	18	0.6	707	13	BM624639	BM624639 170006874
c	667	18	0.6	708	12	BG862149	BG862149 602795716
	668	18	0.6	708	17	BH084250	BH084250 RPCI-24-3
c	669	18	0.6	710	12	BF167713	BF167713 601774423
c	670	18	0.6	710	13	BI850754	BI850754 imageqc_1
c	671	18	0.6	712	10	BE636604	BE636604 rockefell
c	672	18	0.6	713	13	BI823892	BI823892 603039059
	673	18	0.6	716	10	AV707602	AV707602 AV707602
	674	18	0.6	718	12	BF341602	BF341602 602016022
	675	18	0.6	719	10	AW940604	AW940604 GH15883.3
c	676	18	0.6	719	12	BF489097	BF489097 AT24801.5
c	677	18	0.6	720	10	AW220283	AW220283 EST302766
c	678	18	0.6	721	12	BF867608	BF867608 963092H09
	679	18	0.6	722	12	BG758156	BG758156 602712325
c	680	18	0.6	724	13	BM407663	BM407663 EST581990
c	681	18	0.6	725	12	BF182127	BF182127 601804667
c	682	18	0.6	728	12	BF489125	BF489125 AT24834.5
	683	18	0.6	728	17	AG033889	AG033889 Pan trogl
c	684	18	0.6	729	13	BI771203	BI771203 603059213
	685	18	0.6	731	13	BI733584	BI733584 603352467
c	686	18	0.6	733	9	AA717448	AA717448 vu19e06.r
c	687	18	0.6	733	14	BQ575609	BQ575609 UI-H-EZ1-
c	688	18	0.6	735	12	BG173068	BG173068 602335552
	689	18	0.6	736	17	CNS04U5P	AL307366 Tetraodon
c	690	18	0.6	739	9	AA790490	AA790490 vw05b05.r
	691	18	0.6	739	17	AQ688874	AQ688874 nbxb0078B
c	692	18	0.6	740	10	AV764019	AV764019 AV764019
c	693	18	0.6	743	12	BG887487	BG887487 EST513338
	694	18	0.6	744	13	BI309716	BI309716 EST531126
	695	18	0.6	745	9	AJ449216	AJ449216 AJ449216
	696	18	0.6	749	12	BG772601	BG772601 602720706
c	697	18	0.6	751	14	BQ689942	BQ689942 AGENCOURT
c	698	18	0.6	752	11	AK006944	AK006944 Mus muscu
	699	18	0.6	752	12	BE962516	BE962516 601655825
c	700	18	0.6	752	13	BI694625	BI694625 603347848
	701	18	0.6	754	12	BG468654	BG468654 602510258
c	702	18	0.6	755	9	AL537385	AL537385 AL537385
c	703	18	0.6	756	13	BI733912	BI733912 603355623
c	704	18	0.6	756	17	BH674842	BH674842 BOMOF21TF
	705	18	0.6	765	12	BE797787	BE797787 601586388
c	706	18	0.6	767	10	BE377558	BE377558 601229746
c	707	18	0.6	768	12	BF631334	BF631334 HVSMEb001
c	708	18	0.6	768	12	BG887486	BG887486 EST513337
c	709	18	0.6	769	10	BE389523	BE389523 601282874
c	710	18	0.6	769	17	BH466799	BH466799 BOGRC48TR
c	711	18	0.6	770	9	AL518386	AL518386 AL518386
c	712	18	0.6	772	17	CNS046LD	AL276826 Tetraodon
c	713	18	0.6	774	10	BE282936	BE282936 601100409
	714	18	0.6	775	12	BG474100	BG474100 602516671
c	715	18	0.6	776	13	BI652143	BI652143 603299152
c	716	18	0.6	777	12	BG243046	BG243046 602355336
	717	18	0.6	779	10	BE266129	BE266129 601190947
c	718	18	0.6	779	12	BF535768	BF535768 602051273
	719	18	0.6	781	17	CNS077KV	AL432821 T3 end of
c	720	18	0.6	785	12	BG066275	BG066275 H3041A09-
c	721	18	0.6	785	13	BI415864	BI415864 602988072
c	722	18	0.6	786	13	BM411890	BM411890 EST586217
c	723	18	0.6	789	13	BG915330	BG915330 602814655
	724	18	0.6	794	9	AA796623	AA796623 vp31g03.r
c	725	18	0.6	794	12	BE972132	BE972132 601651591
c	726	18	0.6	794	17	BH330657	BH330657 CH230-125
c	727	18	0.6	798	12	BF127237	BF127237 601650894
	728	18	0.6	799	14	BQ953020	BQ953020 AGENCOURT
	729	18	0.6	802	11	AY104906	AY104906 Zea mays
	730	18	0.6	804	12	BF260468	BF260468 HVSMEf002
	731	18	0.6	807	12	BF204907	BF204907 601866720
	732	18	0.6	807	13	BM048518	BM048518 603625032
c	733	18	0.6	808	13	BG914937	BG914937 602813795
c	734	18	0.6	812	13	BI105207	BI105207 602893483
c	735	18	0.6	814	13	BI698032	BI698032 603344182
c	736	18	0.6	816	13	BI150497	BI150497 602915153

c	737	18	0.6	820	11	AK020874	AK020874 Mus muscu
c	738	18	0.6	825	13	BI330469	BI330469 602983484
	739	18	0.6	835	9	AA829751	AA829751 od35e12.s
c	740	18	0.6	836	10	BE378731	BE378731 601237171
c	741	18	0.6	837	12	BG473084	BG473084 602515052
c	742	18	0.6	841	13	BI691860	BI691860 603307584
c	743	18	0.6	845	9	AL559252	AL559252 AL559252
c	744	18	0.6	845	12	BF788104	BF788104 602113314
	745	18	0.6	846	13	BM359463	BM359463 GA__Ea002
	746	18	0.6	849	12	BF666644	BF666644 602121328
	747	18	0.6	852	12	BF314916	BF314916 601899420
c	748	18	0.6	855	14	BQ505020	BQ505020 EST612435
c	749	18	0.6	856	13	BI146833	BI146833 602911729
	750	18	0.6	857	17	AQ747018	AQ747018 HS_5538_A
c	751	18	0.6	860	12	BG328120	BG328120 602427176
	752	18	0.6	860	13	BI854946	BI854946 603381937
	753	18	0.6	862	12	BG723207	BG723207 602690812
c	754	18	0.6	867	13	BI655661	BI655661 603283948
	755	18	0.6	869	12	BF666333	BF666333 602119251
c	756	18	0.6	871	9	AL526108	AL526108 AL526108
	757	18	0.6	872	13	BM019773	BM019773 603648092
c	758	18	0.6	873	9	AL539772	AL539772 AL539772
c	759	18	0.6	873	13	BI414418	BI414418 602986293
c	760	18	0.6	874	13	BI655236	BI655236 603284051
c	761	18	0.6	876	9	AL537290	AL537290 AL537290
	762	18	0.6	876	12	BF217835	BF217835 601882681
	763	18	0.6	878	12	BG035119	BG035119 602324814
c	764	18	0.6	880	12	BF537713	BF537713 602054405
c	765	18	0.6	881	13	BM019895	BM019895 603648276
c	766	18	0.6	882	13	BI853668	BI853668 603380804
c	767	18	0.6	883	13	BI408049	BI408049 602919492
	768	18	0.6	883	17	CNS01HSP	AL144842 Anopheles
c	769	18	0.6	885	17	CNS04K4R	AL294372 Tetraodon
	770	18	0.6	885	17	CNS04NCT	AL298550 Tetraodon
c	771	18	0.6	886	9	AL539785	AL539785 AL539785
	772	18	0.6	886	10	BE542860	BE542860 601067679
c	773	18	0.6	890	12	BF160408	BF160408 601771368
c	774	18	0.6	892	12	BG714423	BG714423 602669952
c	775	18	0.6	895	14	BQ963594	BQ963594 AGENCOURT
c	776	18	0.6	899	14	BQ924538	BQ924538 AGENCOURT
	777	18	0.6	903	17	AZ691717	AZ691717 ENTIS01TR
	778	18	0.6	905	13	BI175496	BI175496 OSTR047E2
c	779	18	0.6	906	14	BQ677208	BQ677208 AGENCOURT
c	780	18	0.6	906	14	BQ960087	BQ960087 AGENCOURT
c	781	18	0.6	907	14	BQ945752	BQ945752 AGENCOURT
c	782	18	0.6	907	17	CNS02PPF	AL208284 Tetraodon
c	783	18	0.6	908	17	CNS02CGW	AL191129 Tetraodon
c	784	18	0.6	909	13	BI904890	BI904890 603168926
	785	18	0.6	909	14	BQ671929	BQ671929 AGENCOURT
	786	18	0.6	912	12	BG116981	BG116981 602319048
	787	18	0.6	915	14	BQ648913	BQ648913 AGENCOURT
c	788	18	0.6	916	14	BQ963508	BQ963508 AGENCOURT
c	789	18	0.6	918	14	BQ425118	BQ425118 AGENCOURT
	790	18	0.6	923	14	BQ900972	BQ900972 AGENCOURT
c	791	18	0.6	924	12	BF301320	BF301320 602029766
	792	18	0.6	926	14	BQ686999	BQ686999 AGENCOURT
c	793	18	0.6	932	12	BE796901	BE796901 601587255
c	794	18	0.6	933	17	BH147131	BH147131 ENTPQ76TR
c	795	18	0.6	941	14	BQ676180	BQ676180 AGENCOURT
c	796	18	0.6	944	17	CNS02JCB	AL200036 Tetraodon
	797	18	0.6	948	12	BF789245	BF789245 602105144
c	798	18	0.6	949	12	BF234946	BF234946 602027390
	799	18	0.6	949	13	BI825673	BI825673 603072655
	800	18	0.6	953	12	BG469969	BG469969 602533441
c	801	18	0.6	955	9	AL558823	AL558823 AL558823
c	802	18	0.6	956	14	BQ964027	BQ964027 AGENCOURT
c	803	18	0.6	959	12	BG863102	BG863102 602798160
c	804	18	0.6	959	14	BQ922851	BQ922851 AGENCOURT
	805	18	0.6	964	13	BI552167	BI552167 603195216
c	806	18	0.6	964	14	BQ942017	BQ942017 AGENCOURT
c	807	18	0.6	966	13	BI656854	BI656854 603284151
c	808	18	0.6	969	12	BG247627	BG247627 602359313
	809	18	0.6	969	12	BG256134	BG256134 602367263

	810	18	0.6	969	17	CNS06W16	AL417856 T3 end of
c	811	18	0.6	971	14	BQ954296	BQ954296 AGENCOURT
	812	18	0.6	974	9	AL551463	AL551463 AL551463
c	813	18	0.6	974	17	CNS03EI0	AL240417 Tetraodon
c	814	18	0.6	977	13	BI411597	BI411597 602965824
c	815	18	0.6	985	14	BQ877325	BQ877325 AGENCOURT
c	816	18	0.6	985	17	CNS01UPO	AL168117 Tetraodon
	817	18	0.6	989	12	BG332198	BG332198 602433528
c	818	18	0.6	993	17	CNS00DT0	AL075286 Drosophil
	819	18	0.6	1004	13	BI551552	BI551552 603194394
	820	18	0.6	1007	17	CNS040DF	AL268764 Tetraodon
c	821	18	0.6	1012	9	AL530671	AL530671 AL530671
c	822	18	0.6	1022	14	BQ964327	BQ964327 AGENCOURT
	823	18	0.6	1024	17	CNS03OPP	AL253654 Tetraodon
c	824	18	0.6	1025	13	BM557677	BM557677 AGENCOURT
c	825	18	0.6	1026	14	BQ709583	BQ709583 AGENCOURT
	826	18	0.6	1030	17	CNS04CBW	AL284261 Tetraodon
c	827	18	0.6	1040	17	CNS03ZQ3	AL267924 Tetraodon
c	828	18	0.6	1045	13	BI658550	BI658550 603283777
	829	18	0.6	1047	13	BM544499	BM544499 AGENCOURT
c	830	18	0.6	1049	12	BF577950	BF577950 602091972
	831	18	0.6	1050	9	AV214969	AV214969 AV214969
c	832	18	0.6	1052	17	CNS00ITV	AL075434 Drosophil
c	833	18	0.6	1065	13	BI411015	BI411015 602963370
	834	18	0.6	1068	14	BM909849	BM909849 AGENCOURT
	835	18	0.6	1071	13	BM560489	BM560489 AGENCOURT
c	836	18	0.6	1075	14	BM925478	BM925478 AGENCOURT
	837	18	0.6	1085	12	BF578325	BF578325 602092683
	838	18	0.6	1088	12	BF209692	BF209692 601874226
c	839	18	0.6	1091	14	BQ944885	BQ944885 AGENCOURT
c	840	18	0.6	1092	17	CNS05MXA	AL344647 Tetraodon
	841	18	0.6	1093	17	CNS03KIQ	AL248219 Tetraodon
	842	18	0.6	1097	13	BG913347	BG913347 602812071
	843	18	0.6	1098	11	AY111914	AY111914 Zea mays
c	844	18	0.6	1098	12	BG248147	BG248147 602400569
c	845	18	0.6	1101	17	CNS008Q0	AL052289 Drosophil
c	846	18	0.6	1101	17	CNS00E3W	AL068709 Drosophil
c	847	18	0.6	1124	14	BQ923373	BQ923373 AGENCOURT
c	848	18	0.6	1129	14	BM910044	BM910044 AGENCOURT
c	849	18	0.6	1130	14	BQ684644	BQ684644 AGENCOURT
c	850	18	0.6	1131	14	BQ276399	BQ276399 AGENCOURT
	851	18	0.6	1131	17	CNS03L5Y	AL249055 Tetraodon
c	852	18	0.6	1147	14	BQ071836	BQ071836 AGENCOURT
	853	18	0.6	1150	14	BQ926132	BQ926132 AGENCOURT
c	854	18	0.6	1179	12	BF181455	BF181455 601805645
	855	18	0.6	1186	17	AG032524	AG032524 Pan trogl
	856	18	0.6	1190	14	BM922179	BM922179 AGENCOURT
c	857	18	0.6	1196	13	BM546078	BM546078 AGENCOURT
c	858	18	0.6	1212	14	BM925803	BM925803 AGENCOURT
c	859	18	0.6	1328	11	AK011107	AK011107 Mus muscu
	860	18	0.6	1337	12	BG471056	BG471056 602512043
c	861	18	0.6	1338	14	BQ684684	BQ684684 AGENCOURT
c	862	18	0.6	1359	14	BM805497	BM805497 AGENCOURT
c	863	18	0.6	1417	14	BM914673	BM914673 AGENCOURT
	864	18	0.6	1434	11	AY106315	AY106315 Zea mays
c	865	18	0.6	1444	14	BM921764	BM921764 AGENCOURT
c	866	18	0.6	1486	17	AG182550	AG182550 Pan trogl
	867	18	0.6	1488	13	BM556903	BM556903 AGENCOURT
c	868	18	0.6	1489	13	BM467281	BM467281 AGENCOURT
c	869	18	0.6	1547	14	BM911629	BM911629 AGENCOURT
	870	18	0.6	1568	13	BM048762	BM048762 603625393
	871	18	0.6	1780	12	BF581621	BF581621 602099973
	872	18	0.6	2420	10	BE402959	BE402959 GBX002.G1
	873	18	0.6	2898	11	AK004737	AK004737 Mus muscu
c	874	17	0.5	54	12	BE783642	BE783642 601471219
	875	17	0.5	80	9	AA761381	AA761381 nz25g12.s
	876	17	0.5	84	9	AA848081	AA848081 od65b11.s
	877	17	0.5	91	9	AA889872	AA889872 al52c06.s
	878	17	0.5	100	9	AA208122	AA208122 mv93d11.r
c	879	17	0.5	100	12	BE842533	BE842533 IL3-ST014
	880	17	0.5	101	9	AA765673	AA765673 oa04f12.s
	881	17	0.5	105	9	AA504849	AA504849 aa64f06.s
c	882	17	0.5	109	12	BE842525	BE842525 IL3-ST014
c	883	17	0.5	109	12	BE842535	BE842535 IL3-ST014
	884	17	0.5	110	12	BF871664	BF871664 QV0-ET015
c	885	17	0.5	111	13	BI727422	BI727422 1031092C0
c	886	17	0.5	112	12	BE842522	BE842522 IL3-ST014
c	887	17	0.5	112	12	BE842524	BE842524 IL3-ST014
c	888	17	0.5	113	12	BE842537	BE842537 IL3-ST014
c	889	17	0.5	114	12	BE842532	BE842532 IL3-ST014
c	890	17	0.5	114	12	BE842534	BE842534 IL3-ST014
c	891	17	0.5	115	12	BE988634	BE988634 UI-M-CG0p
	892	17	0.5	117	14	H92039	H92039 ys82e12.s1
	893	17	0.5	126	14	BQ458388	BQ458388 HA05G04r
c	894	17	0.5	127	12	BE983793	BE983793 UI-M-CG0p
c	895	17	0.5	130	12	BF460938	BF460938 UI-M-CG0p
c	896	17	0.5	132	12	BF464665	BF464665 UI-M-CG0p
c	897	17	0.5	132	12	BF465949	BF465949 UI-M-CG0p
	898	17	0.5	133	9	AA731348	AA731348 nz94d10.s
	899	17	0.5	133	9	AA827636	AA827636 od55g02.s
	900	17	0.5	136	9	AA139263	AA139263 mq02e06.r
c	901	17	0.5	136	9	AV173307	AV173307 AV173307
	902	17	0.5	137	9	AA724434	AA724434 ah91g11.s
	903	17	0.5	140	10	AW197526	AW197526 xm39f06.x
c	904	17	0.5	140	12	BG012637	BG012637 QV3-GN033
c	905	17	0.5	140	17	BH223560	BH223560 1006113F0
c	906	17	0.5	141	9	AA743332	AA743332 ny22f03.s
	907	17	0.5	147	14	BQ636167	BQ636167 hd05d05.y
	908	17	0.5	148	12	BF173209	BF173209 MYE0651a
c	909	17	0.5	149	12	BF817967	BF817967 MR2-CI012
c	910	17	0.5	151	17	BH199133	BH199133 TC3-70M11
	911	17	0.5	152	9	AA231901	AA231901 RZ543.R c
c	912	17	0.5	156	9	AV033363	AV033363 AV033363
c	913	17	0.5	156	14	U44248	U44248 ENU44248 As
c	914	17	0.5	157	12	BE842527	BE842527 IL3-ST014
	915	17	0.5	158	9	AA805512	AA805512 oc12f04.s
c	916	17	0.5	162	9	AL842063	AL842063 AL842063
	917	17	0.5	163	9	AA714181	AA714181 nw05f04.s
c	918	17	0.5	163	9	AI089072	AI089072 ou86g09.s
c	919	17	0.5	165	9	AI842244	AI842244 UI-M-AI1-
	920	17	0.5	165	10	AW603908	AW603908 RC2-CN005
c	921	17	0.5	166	10	AW854556	AW854556 RC3-CT025
	922	17	0.5	166	13	BI908503	BI908503 603069269
	923	17	0.5	166	17	BH224557	BH224557 1006119G0
c	924	17	0.5	169	9	AA928272	AA928272 on79c06.s
	925	17	0.5	170	9	AA014090	AA014090 mh28d03.r
c	926	17	0.5	178	13	BG956065	BG956065 CM4-CT065
c	927	17	0.5	178	17	TA170A12P	AL473303 T. brucei
c	928	17	0.5	180	9	AJ467324	AJ467324 AJ467324
c	929	17	0.5	180	9	AJ468324	AJ468324 AJ468324
c	930	17	0.5	180	17	BH196416	BH196416 TC3-34F10
	931	17	0.5	181	12	BE842526	BE842526 IL3-ST014
c	932	17	0.5	184	14	D22017	D22017 RICC10187A
	933	17	0.5	185	10	BE492632	BE492632 WHE0559_C
	934	17	0.5	186	9	AA767847	AA767847 oa63f05.s
	935	17	0.5	186	9	AV094674	AV094674 AV094674
	936	17	0.5	186	17	AZ824339	AZ824339 2M0098E20
c	937	17	0.5	187	12	BG305870	BG305870 fm13c05.x
	938	17	0.5	188	9	AA040800	AA040800 zf07g02.s
c	939	17	0.5	190	9	AA785083	AA785083 g4g11a1.r
	940	17	0.5	191	9	AU077190	AU077190 AU077190
	941	17	0.5	192	10	AW862068	AW862068 RC3-CT034
	942	17	0.5	193	12	BE842523	BE842523 IL3-ST014
c	943	17	0.5	195	14	BQ694464	BQ694464 1000607 H
	944	17	0.5	198	9	AA888766	AA888766 am32h12.s
c	945	17	0.5	198	9	AV172966	AV172966 AV172966
	946	17	0.5	198	10	BB269372	BB269372 BB269372
c	947	17	0.5	201	13	BI127995	BI127995 G069P46Y
	948	17	0.5	203	9	AA161523	AA161523 MBAFCE9G0
	949	17	0.5	207	9	AI236315	AI236315 EST232877
	950	17	0.5	208	10	BE563499	BE563499 601334980
c	951	17	0.5	208	14	D42622	D42622 D42622 Rice
	952	17	0.5	210	9	AA767490	AA767490 oa90e10.s
c	953	17	0.5	210	9	AV294428	AV294428 AV294428
c	954	17	0.5	210	10	AW114475	AW114475 rs51g07.y
c	955	17	0.5	210	10	BB291607	BB291607 BB291607

```
  956    17   0.5   211  9  AA827866              AA827866 od56f09.s
c 957    17   0.5   211  10  BE268339              BE268339 601124877
c 958    17   0.5   212  10  BB452639              BB452639 BB452639
  959    17   0.5   213  13  BJ102677              BJ102677 BJ102677
  960    17   0.5   214  13  BI430600              BI430600 949057D11
  961    17   0.5   215  9  AI961874              AI961874 wt40c05.x
c 962    17   0.5   215  17  BH197850              BH197850 TC3-68J19
  963    17   0.5   217  10  AV629958              AV629958 AV629958
c 964    17   0.5   217  10  BB423400              BB423400 BB423400
  965    17   0.5   218  12  BG063522              BG063522 H3008C03-
c 966    17   0.5   218  14  C82543                C82543 C82543 rabb
  967    17   0.5   218  14  C83399                C83399 C83399 rabb
c 968    17   0.5   219  10  BB465931              BB465931 BB465931
c 969    17   0.5   219  17  BH230214              BH230214 1006156F1
c 970    17   0.5   221  10  BB325854              BB325854 BB325854
  971    17   0.5   224  10  BB634583              BB634583 BB634583
  972    17   0.5   224  17  BH224524              BH224524 1006119E1
c 973    17   0.5   225  12  BE842539              BE842539 IL3-ST014
  974    17   0.5   226  9  AV244579              AV244579 AV244579
c 975    17   0.5   226  10  BB424005              BB424005 BB424005
c 976    17   0.5   227  9  AL135512              AL135512 DKFZp762J
  977    17   0.5   228  10  BE157121              BE157121 RC3-HT037
c 978    17   0.5   230  10  BB583566              BB583566 BB583566
c 979    17   0.5   231  14  BQ666728              BQ666728 pb46f10.y
  980    17   0.5   232  10  AV357529              AV357529 AV357529
c 981    17   0.5   233  17  BH197594              BH197594 TC3-63L15
c 982    17   0.5   234  10  AW293828              AW293828 UI-H-BI2-
  983    17   0.5   234  12  BG754562              BG754562 602710194
  984    17   0.5   236  9  AA865863              AA865863 oh49g05.s
c 985    17   0.5   237  10  AV323083              AV323083 AV323083
c 986    17   0.5   238  9  AI083456              AI083456 RRAMCA774
  987    17   0.5   238  9  AI443580              AI443580 sa33h06.x
c 988    17   0.5   238  9  AA493583              AA493583 nh01a02.s
c 989    17   0.5   238  9  AA493674              AA493674 nh01f01.s
c 990    17   0.5   238  17  TA339B01Q             AL492571 T. brucei
c 991    17   0.5   239  9  AV078694              AV078694 AV078694
c 992    17   0.5   240  9  AV177983              AV177983 AV177983
c 993    17   0.5   240  17  AZ050380              AZ050380 GSSTc1122
  994    17   0.5   241  9  AA832202              AA832202 oc95b12.s
  995    17   0.5   241  12  BG283562              BG283562 602407876
c 996    17   0.5   242  10  AW500334              AW500334 UI-HF-BN0
  997    17   0.5   243  9  AA913185              AA913185 om94g06.s
  998    17   0.5   243  10  AW865274              AW865274 PM4-SN002
c 999    17   0.5   244  9  AI804116              AI804116 tc68d11.x
c1000    17   0.5   245  9  AU223325              AU223325 AU223325
```

ALIGNMENTS

```
RESULT 1
AQ024974/c
LOCUS       AQ024974                 123 bp    DNA     linear   GSS 23-AUG-2000
DEFINITION  EP(2)0842-5prime Drosophila melanogaster EP line Drosophila
            melanogaster genomic Sequence recovered from 5' end of P element,
            DNA sequence.
ACCESSION   AQ024974
VERSION     AQ024974.1  GI:3265326
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 123)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 116 in the 123 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .123
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT       27 a      28 c      34 g      34 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 17;  Length 123;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     93 GCAGCAACAGCAGAGCAACAACAGC 69


RESULT 2
BH759360/c
LOCUS       BH759360                 365 bp    DNA     linear   GSS 12-MAR-2002
DEFINITION  KG02765 Drosophila melanogaster P{SUPor-P} P element insertion
            lines Drosophila melanogaster genomic Sequence recovered from Both
            5' and 3' ends of P element, DNA sequence.
ACCESSION   BH759360
VERSION     BH759360.1  GI:19352599
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Levis,R., Hoskins,R., Liao,G., Mozden,N., Tsang,G., He,Y., Karpen
            ,G., Bellen,H., Rubin,G. and Spradling,A.
  TITLE     The Berkeley Drosophila Genome Project Gene Disruption Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.
             Sequence orientation is forward strand relative to 5' end of P
            element
             The P element insertion position is base 171 in the 365 bases.
            This insertion position refers to the first base of the 8 base
            target recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
```

```
     source          1..365
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster P{SUPor-P} P element
                     insertion lines"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains one or more
                     P{SUPor-P} P-element transposon insertion. The resultant
                     fragment for each strain was directly sequenced to
                     determine the genomic sequence at the site of insertion.
                     Details of the protocols used can be found at
                     http://www.fruitfly.org/about/methods/inverse.pcr.html."
BASE COUNT       62 a     84 c    101 g    118 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 17;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db        153 GCAGCAACAGCAGAGCAACAACAGC 129


RESULT 3
BI241907
LOCUS       BI241907                 431 bp    mRNA    linear   EST 12-JUL-2001
DEFINITION  RE39108.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE39108 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/12/2001,
            mRNA sequence.
ACCESSION   BI241907
VERSION     BI241907.1  GI:14710636
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/12/2001
            Plate: RE.391  row: A  column: 8
            High quality sequence stop: 430.
FEATURES             Location/Qualifiers
     source          1..431
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE39108"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      125 a    100 c     87 g    119 t
```

```
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 431;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 4
BI482365
LOCUS       BI482365                 440 bp    mRNA    linear   EST 27-AUG-2001
DEFINITION  RE65049.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE65049 5 similar to CG3036:
            FBan0003036 GO:[transporter (GO:0005215)] located on: 2L 25B5-25B6
            ;: 05/16/2001, mRNA sequence.
ACCESSION   BI482365
VERSION     BI482365.1  GI:15317377
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/16/2001
            Plate: RE.650  row: E  column: 1
            High quality sequence stop: 438.
FEATURES             Location/Qualifiers
     source          1..440
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE65049"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      130 a     99 c     88 g    123 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 440;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         31 GCAGCAACAGCAGAGCAACAACAGC 55


RESULT 5
AI063687
```

```
LOCUS       AI063687                 443 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH03640.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH03640 5prime, mRNA sequence.
ACCESSION   AI063687
VERSION     AI063687.1  GI:3339631
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 36  row: D  column: 4
            High quality sequence stop: 344.
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH03640"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      137 a    100 c     85 g    121 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 443;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         11 GCAGCAACAGCAGAGCAACAACAGC 35


RESULT 6
BI365464
LOCUS       BI365464                 485 bp    mRNA    linear   EST 01-AUG-2001
DEFINITION  RE50784.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE50784 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/14/2001,
            mRNA sequence.
ACCESSION   BI365464
VERSION     BI365464.1  GI:15061492
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/14/2001
            Plate: RE.507  row: G  column: 12
            High quality sequence stop: 420.
FEATURES             Location/Qualifiers
     source          1..485
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE50784"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      143 a    109 c     97 g    136 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 485;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 7
BI165260
LOCUS       BI165260                 491 bp    mRNA    linear   EST 09-JUL-2001
DEFINITION  RE05115.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE05115 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 04/11/2001,
            mRNA sequence.
ACCESSION   BI165260
VERSION     BI165260.1  GI:14631066
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 491)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 04/11/2001
            Plate: RE.51  row: B  column: 3
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1..491
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
```

```
                     /clone="RE05115"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      146 a    110 c     96 g    139 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 13;  Length 491;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 8
BI227520
LOCUS       BI227520                 521 bp    mRNA    linear   EST 11-JUL-2001
DEFINITION  RE24591.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE24591 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 04/12/2001,
            mRNA sequence.
ACCESSION   BI227520
VERSION     BI227520.1  GI:14694784
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 521)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 04/12/2001
            Plate: RE.245  row: H  column: 7
            High quality sequence stop: 445.
FEATURES             Location/Qualifiers
     source          1. .521
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE24591"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      154 a    119 c    104 g    144 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 13;  Length 521;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 9
BI363907
LOCUS       BI363907                 526 bp    mRNA    linear   EST 01-AUG-2001
DEFINITION  RE48884.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE48884 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/13/2001,
            mRNA sequence.
ACCESSION   BI363907
VERSION     BI363907.1  GI:15059935
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 526)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/13/2001
            Plate: RE.488  row: G  column: 12
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1. .526
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE48884"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      157 a    117 c    108 g    144 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 13;  Length 526;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 10
BI356933
LOCUS       BI356933                 531 bp    mRNA    linear   EST 31-JUL-2001
DEFINITION  RE43234.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
```

```
            Drosophila melanogaster cDNA clone RE43234 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/13/2001,
            mRNA sequence.
ACCESSION   BI356933
VERSION     BI356933.1  GI:15051387
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/13/2001
            Plate: RE.432  row: C  column: 10
            High quality sequence stop: 486.
FEATURES             Location/Qualifiers
     source          1..531
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE43234"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      158 a    119 c    107 g    147 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 531;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     35 GCAGCAACAGCAGAGCAACAACAGC 59


RESULT 11
AI114180
LOCUS       AI114180                 544 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH10888.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH10888 5prime, mRNA sequence.
ACCESSION   AI114180
VERSION     AI114180.1  GI:3514983
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 108  row: H  column: 4
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
     source          1..544
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH10888"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      163 a    120 c    116 g    145 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     12 GCAGCAACAGCAGAGCAACAACAGC 36


RESULT 12
AI135320
LOCUS       AI135320                 545 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH13005.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH13005 5prime, mRNA sequence.
ACCESSION   AI135320
VERSION     AI135320.1  GI:3627878
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 130  row: A  column: 5
            High quality sequence stop: 544.
FEATURES             Location/Qualifiers
     source          1..545
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH13005"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      163 a    121 c    114 g    147 t
ORIGIN
```

```
  Query Match             0.8%;  Score 25;  DB 9;  Length 545;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      15 GCAGCAACAGCAGAGCAACAACAGC 39


RESULT 13
BI362358
LOCUS       BI362358                 560 bp    mRNA    linear   EST 01-AUG-2001
DEFINITION  RE46895.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE46895 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/13/2001,
            mRNA sequence.
ACCESSION   BI362358
VERSION     BI362358.1  GI:15058386
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 560)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/13/2001
            Plate: RE.468  row: H  column: 11
            High quality sequence stop: 497.
FEATURES             Location/Qualifiers
     source          1..560
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE46895"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      167 a    123 c    117 g    153 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 560;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      31 GCAGCAACAGCAGAGCAACAACAGC 55


RESULT 14
BI164055
LOCUS       BI164055                 577 bp    mRNA    linear   EST 09-JUL-2001
DEFINITION  RE03625.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE03625 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 04/11/2001,
            mRNA sequence.
ACCESSION   BI164055
VERSION     BI164055.1  GI:14629861
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 04/11/2001
            Plate: RE.36  row: C  column: 1
            High quality sequence stop: 497.
FEATURES             Location/Qualifiers
     source          1..577
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE03625"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      168 a    131 c    120 g    158 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 577;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 15
BI165486
LOCUS       BI165486                 579 bp    mRNA    linear   EST 09-JUL-2001
DEFINITION  RE05402.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE05402 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 04/11/2001,
            mRNA sequence.
ACCESSION   BI165486
VERSION     BI165486.1  GI:14631292
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 579)
```

```
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 04/11/2001
            Plate: RE.54  row: A  column: 2
            High quality sequence stop: 561.
FEATURES             Location/Qualifiers
     source          1..579
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE05402"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      170 a    131 c    122 g    156 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      143 GCAGCAACAGCAGAGCAACAACAGC 167
            |||||||||||||||||||||||||
Db       33 GCAGCAACAGCAGAGCAACAACAGC 57


RESULT 16
AQ074074
LOCUS       AQ074074                 583 bp    DNA     linear   GSS 23-AUG-2000
DEFINITION  EP(2)2068 Drosophila melanogaster EP line Drosophila melanogaster
            genomic Sequence recovered from Both 5' and 3' ends of P element,
            DNA sequence.
ACCESSION   AQ074074
VERSION     AQ074074.1  GI:3403324
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 155 in the 583 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1..583
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT      210 a    148 c    130 g     95 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 17;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      143 GCAGCAACAGCAGAGCAACAACAGC 167
            |||||||||||||||||||||||||
Db      491 GCAGCAACAGCAGAGCAACAACAGC 515


RESULT 17
AQ073492/c
LOCUS       AQ073492                 586 bp    DNA     linear   GSS 23-AUG-2000
DEFINITION  EP(2)2323 Drosophila melanogaster EP line Drosophila melanogaster
            genomic Sequence recovered from Both 5' and 3' ends of P element,
            DNA sequence.
ACCESSION   AQ073492
VERSION     AQ073492.1  GI:3403534
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 586)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 116 in the 586 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1..586
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
```

```
                /clone_lib="Drosophila melanogaster EP line"
                /note="Inverse PCR was performed on Drosophila
                melanogaster strains each of which contains a single EP
                transposable element insertion. (The generation of these
                insertion strains is described in Rorth P, Szabo K, Bailey
                A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                genetics in Drosophila. Development 6:1049-1057.) The
                resultant fragment for each strain was directly sequenced
                to determine the genomic sequence at the site of
                insertion. Details of the protocols used can be found at
                http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT       96 a    131 c    151 g    208 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 17;  Length 586;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      93 GCAGCAACAGCAGAGCAACAACAGC 69


RESULT 18
AI387965
LOCUS       AI387965                 590 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH18695.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH18695 5prime, mRNA sequence.
ACCESSION   AI387965
VERSION     AI387965.1  GI:4201987
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 186  row: H  column: 11
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1..590
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH18695"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      172 a    135 c    125 g    158 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 9;  Length 590;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      15 GCAGCAACAGCAGAGCAACAACAGC 39


RESULT 19
BI238617
LOCUS       BI238617                 590 bp    mRNA    linear   EST 12-JUL-2001
DEFINITION  RE35121.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE35121 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/12/2001,
            mRNA sequence.
ACCESSION   BI238617
VERSION     BI238617.1  GI:14707119
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/12/2001
            Plate: RE.351  row: B  column: 9
            High quality sequence stop: 491.
FEATURES             Location/Qualifiers
     source          1..590
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE35121"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      173 a    135 c    123 g    159 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 13;  Length 590;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 20
BI363426
LOCUS       BI363426                 592 bp    mRNA    linear   EST 01-AUG-2001
DEFINITION  RE48231.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE48231 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/13/2001,
            mRNA sequence.
ACCESSION   BI363426
VERSION     BI363426.1  GI:15059454
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
```

```
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/13/2001
            Plate: RE.482  row: C  column: 7
            High quality sequence stop: 472.
FEATURES             Location/Qualifiers
     source          1..592
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE48231"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      171 a    137 c    124 g    159 t      1 others
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 21
BI481998
LOCUS       BI481998                 593 bp    mRNA    linear   EST 28-AUG-2001
DEFINITION  RE64636.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE64636 5 similar to CG3036:
            FBan0003036 GO:[transporter (GO:0005215)] located on: 2L 25B5-25B6
            ;: 05/18/2001, mRNA sequence.
ACCESSION   BI481998
VERSION     BI481998.1  GI:15321162
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/18/2001
            Plate: RE.646  row: C  column: 12
            High quality sequence stop: 561.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE64636"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      173 a    136 c    124 g    160 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 593;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 22
AI109244
LOCUS       AI109244                 602 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH08375.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH08375 5prime, mRNA sequence.
ACCESSION   AI109244
VERSION     AI109244.1  GI:3477568
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 83  row: G  column: 3
            High quality sequence stop: 541.
FEATURES             Location/Qualifiers
     source          1..602
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH08375"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
```

```
BASE COUNT      174 a    140 c    129 g    159 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 602;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     12 GCAGCAACAGCAGAGCAACAACAGC 36


RESULT 23
AI107819
LOCUS       AI107819                 616 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH05674.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH05674 5prime, mRNA sequence.
ACCESSION   AI107819
VERSION     AI107819.1  GI:3475472
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 56  row: G  column: 2
            High quality sequence stop: 547.
FEATURES             Location/Qualifiers
     source          1..616
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH05674"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      177 a    146 c    129 g    164 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 616;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     17 GCAGCAACAGCAGAGCAACAACAGC 41


RESULT 24
AI063205
LOCUS       AI063205                 625 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH02745.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH02745 5prime, mRNA sequence.
ACCESSION   AI063205
VERSION     AI063205.1  GI:3339044
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 625)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 27  row: D  column: 9
            High quality sequence stop: 546.
FEATURES             Location/Qualifiers
     source          1..625
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH02745"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      179 a    151 c    129 g    166 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 625;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     17 GCAGCAACAGCAGAGCAACAACAGC 41


RESULT 25
AI238498
LOCUS       AI238498                 625 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH14550.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH14550 5prime, mRNA sequence.
ACCESSION   AI238498
VERSION     AI238498.1  GI:3833356
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 625)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 145  row: E  column: 2
            High quality sequence stop: 518.
FEATURES             Location/Qualifiers
     source          1..625
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH14550"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
```

```
                  /lab_host="DH5 - alpha"
                  /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                  XhoI; Sized fractionated cDNAs were directly ligated into
                  pOT2.  Plasmid cDNA library."
BASE COUNT      179 a    150 c    132 g    164 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 625;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      143 GCAGCAACAGCAGAGCAACAACAGC 167
            |||||||||||||||||||||||||
Db       19 GCAGCAACAGCAGAGCAACAACAGC 43


RESULT 26
BI170514
LOCUS       BI170514                 626 bp    mRNA    linear   EST 09-JUL-2001
DEFINITION  RE12003.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE12003 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 04/12/2001,
            mRNA sequence.
ACCESSION   BI170514
VERSION     BI170514.1  GI:14636321
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 626)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 04/12/2001
            Plate: RE.120  row: A  column: 3
            High quality sequence stop: 554.
FEATURES             Location/Qualifiers
     source          1. .626
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE12003"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      180 a    146 c    132 g    167 t      1 others
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 626;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      143 GCAGCAACAGCAGAGCAACAACAGC 167
            |||||||||||||||||||||||||
Db       32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 27
AI296017
LOCUS       AI296017                 628 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  LP09807.5prime LP Drosophila melanogaster larval-early pupal pOT2
            Drosophila melanogaster cDNA clone LP09807 5prime, mRNA sequence.
ACCESSION   AI296017
VERSION     AI296017.1  GI:3945424
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 628)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 98  row: A  column: 7
            High quality sequence stop: 500.
FEATURES             Location/Qualifiers
     source          1. .628
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="LP09807"
                     /clone_lib="LP Drosophila melanogaster larval-early pupal
                     pOT2"
                     /sex="male and female"
                     /dev_stage="larvae-pupae"
                     /lab_host="DH5-alpha"
                     /note="Organ: whole body; Vector: pOT2; Site_1: EcoRI;
                     Site_2: Xhol; Sized fractionated cDNAs were directly
                     ligated into pOT2. Plasmid cDNA library.
                     "
BASE COUNT      179 a    149 c    134 g    166 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 628;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      143 GCAGCAACAGCAGAGCAACAACAGC 167
            |||||||||||||||||||||||||
Db        3 GCAGCAACAGCAGAGCAACAACAGC 27


RESULT 28
BI169402
LOCUS       BI169402                 633 bp    mRNA    linear   EST 09-JUL-2001
DEFINITION  RE10566.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE10566 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 04/11/2001,
            mRNA sequence.
ACCESSION   BI169402
VERSION     BI169402.1  GI:14635209
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
```

```
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 04/11/2001
            Plate: RE.105  row: F  column: 6
            High quality sequence stop: 560.
FEATURES             Location/Qualifiers
     source          1. .633
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE10566"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      182 a    149 c    132 g    169 t      1 others
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 29
AI386994
LOCUS       AI386994                 639 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH17396.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH17396 5prime, mRNA sequence.
ACCESSION   AI386994
VERSION     AI386994.1  GI:4201005
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 639)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 173  row: H  column: 12
            High quality sequence stop: 546.
FEATURES             Location/Qualifiers
     source          1. .639
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH17396"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      181 a    153 c    136 g    169 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 639;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      17 GCAGCAACAGCAGAGCAACAACAGC 41


RESULT 30
AI388097
LOCUS       AI388097                 655 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH18879.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH18879 5prime, mRNA sequence.
ACCESSION   AI388097
VERSION     AI388097.1  GI:4202108
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 188  row: G  column: 7
            High quality sequence stop: 536.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH18879"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      187 a    159 c    138 g    171 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 655;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      15 GCAGCAACAGCAGAGCAACAACAGC 39


RESULT 31
AI107422
LOCUS       AI107422                 664 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH05102.5prime GH Drosophila melanogaster head pOT2 Drosophila
```

```
            melanogaster cDNA clone GH05102 5prime, mRNA sequence.
ACCESSION   AI107422
VERSION     AI107422.1  GI:3475075
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: GH05102.3prime
            Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 51  row: A  column: 2
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1..664
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH05102"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      189 a    161 c    138 g    175 t      1 others
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     25 GCAGCAACAGCAGAGCAACAACAGC 49


RESULT 32
AI292862
LOCUS       AI292862               664 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH15875.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH15875 5prime, mRNA sequence.
ACCESSION   AI292862
VERSION     AI292862.1  GI:3942269
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 158  row: G  column: 3
            High quality sequence stop: 509.
FEATURES             Location/Qualifiers
     source          1..664
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH15875"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      189 a    160 c    141 g    174 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     17 GCAGCAACAGCAGAGCAACAACAGC 41


RESULT 33
AI109291
LOCUS       AI109291               666 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH08434.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH08434 5prime, mRNA sequence.
ACCESSION   AI109291
VERSION     AI109291.1  GI:3477615
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 84  row: C  column: 10
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1..666
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH08434"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      189 a    162 c    140 g    175 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 666;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     25 GCAGCAACAGCAGAGCAACAACAGC 49
```

```
RESULT 34
BI368317
LOCUS       BI368317               677 bp    mRNA    linear   EST 01-AUG-2001
DEFINITION  RE54268.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE54268 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/14/2001,
            mRNA sequence.
ACCESSION   BI368317
VERSION     BI368317.1  GI:15064345
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/14/2001
            Plate: RE.542  row: F  column: 8
            High quality sequence stop: 559.
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE54268"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      190 a    163 c    145 g    179 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 13;  Length 677;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      30 GCAGCAACAGCAGAGCAACAACAGC 54

RESULT 35
BI374631
LOCUS       BI374631               686 bp    mRNA    linear   EST 01-AUG-2001
DEFINITION  RE62269.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE62269 5 similar to CG3036:
            FBan0003036 GO:[transporter (GO:0005215)] located on: 2L 25B5-25B6
            ;: 05/16/2001, mRNA sequence.
ACCESSION   BI374631
VERSION     BI374631.1  GI:15070659
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/16/2001
            Plate: RE.622  row: F  column: 9
            High quality sequence stop: 637.
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE62269"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      194 a    165 c    146 g    181 t
ORIGIN

  Query Match              0.8%;  Score 25;  DB 13;  Length 686;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     143 GCAGCAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||||||
Db      30 GCAGCAACAGCAGAGCAACAACAGC 54

RESULT 36
BI368263
LOCUS       BI368263               687 bp    mRNA    linear   EST 01-AUG-2001
DEFINITION  RE54204.5prime RE Drosophila melanogaster normalized Embryo pFlc-1
            Drosophila melanogaster cDNA clone RE54204 5 similar to CG3036:
            FBan0003036 'transporter' located on: 2L 25B5-25B6;: 05/14/2001,
            mRNA sequence.
ACCESSION   BI368263
VERSION     BI368263.1  GI:15064291
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
```

```
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003575: arm:2L [4533649,4860857]
            estimated-cyto:25A2-25B8: 05/14/2001
            Plate: RE.542  row: A  column: 4
            High quality sequence stop: 560.
FEATURES             Location/Qualifiers
     source          1..687
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE54204"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT     194 a    167 c    145 g    181 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 13;  Length 687;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         32 GCAGCAACAGCAGAGCAACAACAGC 56


RESULT 37
AI517363
LOCUS       AI517363                 734 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH28085.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH28085 5prime, mRNA sequence.
ACCESSION   AI517363
VERSION     AI517363.1  GI:4420463
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 734)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 280  row: H  column: 1
            High quality sequence stop: 541.
FEATURES             Location/Qualifiers
     source          1..734
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH28085"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT     206 a    176 c    168 g    184 t
ORIGIN

  Query Match             0.8%;  Score 25;  DB 9;  Length 734;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 GCAGCAACAGCAGAGCAACAACAGC 167
              |||||||||||||||||||||||||
Db         12 GCAGCAACAGCAGAGCAACAACAGC 36


RESULT 38
AZ628307/c
LOCUS       AZ628307                 496 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0480O18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0480O18 F, DNA sequence.
ACCESSION   AZ628307
VERSION     AZ628307.1  GI:11750497
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0480  row: O  column: 18
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 496.
FEATURES             Location/Qualifiers
     source          1..496
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0480O18"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
```

```
BASE COUNT       79 a    194 c     70 g    153 t
ORIGIN

  Query Match             0.7%;  Score 21;  DB 17;  Length 496;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2454 CGGTGTTGAAGGCAAAATGCG 2474
           |||||||||||||||||||||
Db      28 CGGTGTTGAAGGCAAAATGCG 8


RESULT 39
BI468065/c
LOCUS       BI468065                 593 bp    mRNA    linear   EST 22-AUG-2001
DEFINITION  EST00475 Atlantic salmon Lambda Zap Express liver cDNA library
            Salmo salar cDNA clone LRR5-g04 5' similar to Hemopexin-like
            protein variant 1, mRNA sequence.
ACCESSION   BI468065
VERSION     BI468065.1  GI:15280943
KEYWORDS    EST.
SOURCE      Atlantic salmon.
  ORGANISM  Salmo salar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Martin, S.A., Caplice, N.C., Davey, G.C. and Powell, R.
  TITLE     An expressed sequence tag based list of genes expressed in the
            liver of Atlantic salmon (Salmo salar)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Martin SA
            Department of Microbiology
            National University of Ireland Galway (NUIG)
            National University of Ireland Galway, Galway, Ireland
            Tel: 00353 91 524411 2254
            Fax: 00353 91 525700
            Email: sarah.martin@nuigalway.ie
            Insert Length: 593   Std Error: 0.00
            Plate: Liver rare plate 5  row: g  column: 04
            Seq primer: M13 reverse primer = caggaaacagctatgacc
            High quality sequence stop: 593
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Salmo salar"
                     /db_xref="taxon:8030"
                     /clone="LRR5-g04"
                     /clone_lib="Atlantic salmon Lambda Zap Express liver cDNA
                     library"
                     /tissue_type="liver"
                     /dev_stage="adult"
                     /note="Organ: liver; Vector: Lambda ZAP Express; Site_1:
                     EcoR1; Site_2: Xho1; An Atlantic salmon liver cDNA library
                     was constructed using the  Lambda Zap Express/Gigapack
                     cloning kit (Stratagene cloning systems). cDNA synthesis
                     was carried out using an oligo-(dT) primer for the
                     reverse transcription of 5ug of mRNA and the library was
                     constructed by  directional cloning EcoR1-Xho1 based on
                     manufacturers instructions. An  insert:vector ligation
                     ratio of 1:5 was chosen as most optimum . The  lambda
                     library was packaged with Gigapack III gold packaging
                     extracts  and plated on the E. coli cell line XL1-Blue
                     MRF'."
BASE COUNT      152 a    182 c    141 g    118 t
ORIGIN

  Query Match             0.7%;  Score 21;  DB 13;  Length 593;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2710 CATGCGGAAGGCAGCGTCAAC 2730
           |||||||||||||||||||||
Db     352 CATGCGGAAGGCAGCGTCAAC 332


RESULT 40
AI238153
LOCUS       AI238153                 727 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  GH14154.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH14154 5prime, mRNA sequence.
ACCESSION   AI238153
VERSION     AI238153.1  GI:3833011
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 727)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 141  row: E  column: 6
            High quality sequence stop: 597.
FEATURES             Location/Qualifiers
     source          1..727
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH14154"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      201 a    173 c    164 g    189 t
ORIGIN

  Query Match             0.7%;  Score 21;  DB 9;  Length 727;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     147 CAACAGCAGAGCAACAACAGC 167
           |||||||||||||||||||||
Db       1 CAACAGCAGAGCAACAACAGC 21


Search completed: January 27, 2003, 23:05:37
Job time : 2950 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 20:57:34 ; Search time 81 Seconds
(without alignments)
12130.770 Million cell updates/sec

Title: US-09-830-433A-7
Perfect score: 3204
Sequence: 1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	20	0.6	735	3	US-08-750-145A-23	Sequence 23, Appl
c	2	20	0.6	735	3	US-08-975-698A-27	Sequence 27, Appl
c	3	20	0.6	735	4	US-09-417-090-27	Sequence 27, Appl
c	4	20	0.6	735	4	US-09-727-578-27	Sequence 27, Appl
c	5	18	0.6	60	1	US-08-219-012-74	Sequence 74, Appl
c	6	18	0.6	60	4	US-08-687-421-262	Sequence 262, App
c	7	18	0.6	692	4	US-09-221-017B-718	Sequence 718, App
c	8	18	0.6	1689	4	US-09-276-531-101	Sequence 101, App
	9	17	0.5	77	1	US-08-248-474-87	Sequence 87, Appl
	10	17	0.5	77	3	US-08-756-849-87	Sequence 87, Appl
	11	17	0.5	117	4	US-08-906-156A-27	Sequence 27, Appl
	12	17	0.5	322	4	US-08-906-156A-28	Sequence 28, Appl
	13	17	0.5	362	4	US-09-641-638-222	Sequence 222, App
	14	17	0.5	362	4	US-09-641-638-223	Sequence 223, App
	15	17	0.5	362	4	US-09-641-638-224	Sequence 224, App
	16	17	0.5	362	4	US-09-641-638-225	Sequence 225, App
	17	17	0.5	429	4	US-09-060-756-653	Sequence 653, App
	18	17	0.5	500	4	US-08-642-274D-18	Sequence 18, Appl
	19	17	0.5	500	4	US-08-952-014C-18	Sequence 18, Appl
	20	17	0.5	671	4	US-08-998-416-826	Sequence 826, App
	21	17	0.5	800	5	PCT-US95-04801-4	Sequence 4, Appli
	22	17	0.5	871	4	US-08-906-156A-13	Sequence 13, Appl
	23	17	0.5	871	4	US-08-906-156A-58	Sequence 58, Appl
	24	17	0.5	925	3	US-09-267-031-5	Sequence 5, Appli
c	25	17	0.5	981	2	US-08-770-565-1	Sequence 1, Appli
c	26	17	0.5	981	2	US-08-710-249-5	Sequence 5, Appli
c	27	17	0.5	981	2	US-08-833-377-1	Sequence 1, Appli
c	28	17	0.5	981	2	US-08-714-482-2	Sequence 2, Appli
c	29	17	0.5	981	3	US-08-838-545-22	Sequence 22, Appl
c	30	17	0.5	981	4	US-09-349-532-22	Sequence 22, Appl
c	31	17	0.5	981	4	US-09-220-157A-5	Sequence 5, Appli
c	32	17	0.5	1035	1	US-07-601-094-30	Sequence 30, Appl
c	33	17	0.5	1035	1	US-08-012-735-30	Sequence 30, Appl
	34	17	0.5	1257	4	US-08-791-115B-8	Sequence 8, Appli
c	35	17	0.5	1440	2	US-08-224-482-5	Sequence 5, Appli
	36	17	0.5	1773	4	US-08-906-156A-11	Sequence 11, Appl
c	37	17	0.5	1773	4	US-09-310-363C-7	Sequence 7, Appli
	38	17	0.5	1801	5	PCT-US95-02455-1	Sequence 1, Appli
c	39	17	0.5	1910	4	US-09-593-711A-3	Sequence 3, Appli
c	40	17	0.5	1914	1	US-07-601-094-1	Sequence 1, Appli
c	41	17	0.5	1914	1	US-08-012-735-1	Sequence 1, Appli
	42	17	0.5	1962	4	US-08-791-115B-3	Sequence 3, Appli
	43	17	0.5	1971	6	5349058-1	Patent No. 5349058
	44	17	0.5	2139	3	US-09-026-343-1	Sequence 1, Appli
	45	17	0.5	2139	4	US-09-362-871-1	Sequence 1, Appli
c	46	17	0.5	2406	1	US-08-396-479B-7	Sequence 7, Appli
c	47	17	0.5	2406	1	US-08-818-823-7	Sequence 7, Appli
c	48	17	0.5	2420	1	US-08-330-123A-3	Sequence 3, Appli
c	49	17	0.5	2420	4	US-09-580-517-3	Sequence 3, Appli
c	50	17	0.5	2425	2	US-08-485-778-1	Sequence 1, Appli
c	51	17	0.5	2425	3	US-08-520-550A-1	Sequence 1, Appli
c	52	17	0.5	2426	1	US-08-482-115B-3	Sequence 3, Appli
c	53	17	0.5	2426	2	US-08-660-678A-3	Sequence 3, Appli
c	54	17	0.5	2426	2	US-08-472-802C-4	Sequence 4, Appli
c	55	17	0.5	2426	2	US-08-714-482-1	Sequence 1, Appli
c	56	17	0.5	2426	3	US-08-998-443-3	Sequence 3, Appli
c	57	17	0.5	2426	4	US-09-060-523-3	Sequence 3, Appli
c	58	17	0.5	2647	1	US-08-396-479B-9	Sequence 9, Appli
c	59	17	0.5	2647	1	US-08-818-823-9	Sequence 9, Appli
c	60	17	0.5	2759	4	US-09-144-367-1	Sequence 1, Appli
	61	17	0.5	2796	2	US-08-937-931-3	Sequence 3, Appli
	62	17	0.5	2796	4	US-09-285-502-3	Sequence 3, Appli
	63	17	0.5	2796	4	US-09-709-126-3	Sequence 3, Appli
	64	17	0.5	2796	4	US-09-871-385A-3	Sequence 3, Appli
c	65	17	0.5	2811	1	US-08-040-548-31	Sequence 31, Appl
c	66	17	0.5	2811	1	US-08-466-344-31	Sequence 31, Appl
c	67	17	0.5	2817	6	5206152-6	Patent No. 5206152
	68	17	0.5	3160	4	US-08-791-115B-2	Sequence 2, Appli
	69	17	0.5	3330	1	US-08-149-103-1	Sequence 1, Appli
	70	17	0.5	3330	1	US-08-451-883-1	Sequence 1, Appli
	71	17	0.5	3349	2	US-08-920-234-1	Sequence 1, Appli
	72	17	0.5	3410	4	US-09-527-154-3	Sequence 3, Appli
	73	17	0.5	3656	1	US-08-393-734-1	Sequence 1, Appli
	74	17	0.5	3656	4	US-08-894-489-1	Sequence 1, Appli
	75	17	0.5	3695	4	US-09-453-702B-18	Sequence 18, Appl
c	76	17	0.5	3969	1	US-08-396-479B-11	Sequence 11, Appl
c	77	17	0.5	3969	1	US-08-818-823-11	Sequence 11, Appl
c	78	17	0.5	4010	5	PCT-US94-07297-40	Sequence 40, Appl
c	79	17	0.5	4534	2	US-08-935-450-1	Sequence 1, Appli
c	80	17	0.5	4534	4	US-09-166-350-25	Sequence 25, Appl
	81	17	0.5	8147	4	US-09-514-247A-9	Sequence 9, Appli
	82	17	0.5	9592	1	US-08-393-734-3	Sequence 3, Appli
	83	17	0.5	9592	4	US-08-894-489-3	Sequence 3, Appli
	84	17	0.5	10747	2	US-08-147-777-1	Sequence 1, Appli
	85	17	0.5	10747	3	US-08-452-872-1	Sequence 1, Appli
	86	17	0.5	10747	5	PCT-US93-03985-1	Sequence 1, Appli
	87	17	0.5	24979	2	US-08-147-777-3	Sequence 3, Appli
	88	17	0.5	24979	3	US-08-452-872-3	Sequence 3, Appli
	89	17	0.5	24979	5	PCT-US93-03985-3	Sequence 3, Appli
	90	17	0.5	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	91	17	0.5	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	92	17	0.5	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	93	17	0.5	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	94	16	0.5	18	2	US-09-205-860-10	Sequence 10, Appl
	95	16	0.5	20	4	US-09-517-584A-13	Sequence 13, Appl
c	96	16	0.5	322	4	US-09-117-121-41	Sequence 41, Appl
c	97	16	0.5	322	4	US-09-344-529-7	Sequence 7, Appli
	98	16	0.5	386	3	US-08-804-439A-96	Sequence 96, Appl
	99	16	0.5	386	3	US-08-720-229-96	Sequence 96, Appl
	100	16	0.5	390	4	US-08-930-264-3	Sequence 3, Appli

	101	16	0.5	390	4	US-08-930-264-5	Sequence 5, Appli
	102	16	0.5	390	4	US-08-930-264-7	Sequence 7, Appli
	103	16	0.5	390	4	US-08-930-264-9	Sequence 9, Appli
	104	16	0.5	390	4	US-08-930-264-11	Sequence 11, Appl
	105	16	0.5	390	4	US-08-930-264-13	Sequence 13, Appl
	106	16	0.5	390	4	US-08-930-264-15	Sequence 15, Appl
	107	16	0.5	390	4	US-08-930-264-17	Sequence 17, Appl
	108	16	0.5	390	4	US-08-930-264-19	Sequence 19, Appl
	109	16	0.5	390	4	US-08-930-264-21	Sequence 21, Appl
	110	16	0.5	390	4	US-08-930-264-23	Sequence 23, Appl
	111	16	0.5	390	4	US-08-930-264-25	Sequence 25, Appl
	112	16	0.5	416	4	US-09-370-838-249	Sequence 249, App
	113	16	0.5	440	4	US-09-397-787-328	Sequence 328, App
c	114	16	0.5	451	4	US-09-357-251-3	Sequence 3, Appli
	115	16	0.5	457	2	US-08-332-766A-7	Sequence 7, Appli
	116	16	0.5	491	1	US-07-945-288-7	Sequence 7, Appli
	117	16	0.5	491	1	US-08-462-831-7	Sequence 7, Appli
	118	16	0.5	491	1	US-08-461-809-7	Sequence 7, Appli
	119	16	0.5	491	1	US-08-461-441-7	Sequence 7, Appli
	120	16	0.5	491	1	US-08-482-142-7	Sequence 7, Appli
	121	16	0.5	491	2	US-08-478-572-7	Sequence 7, Appli
	122	16	0.5	491	4	US-08-484-296-7	Sequence 7, Appli
c	123	16	0.5	491	4	US-09-643-597-307	Sequence 307, App
	124	16	0.5	491	5	PCT-US93-08518-7	Sequence 7, Appli
c	125	16	0.5	502	4	US-08-893-654B-7	Sequence 7, Appli
	126	16	0.5	511	4	US-08-936-165A-182	Sequence 182, App
	127	16	0.5	513	1	US-08-288-888-3	Sequence 3, Appli
	128	16	0.5	513	2	US-08-910-075-3	Sequence 3, Appli
	129	16	0.5	513	2	US-08-905-801A-3	Sequence 3, Appli
	130	16	0.5	516	1	US-08-532-828B-13	Sequence 13, Appl
	131	16	0.5	516	1	US-08-532-828B-14	Sequence 14, Appl
	132	16	0.5	588	4	US-09-565-808-1	Sequence 1, Appli
	133	16	0.5	663	2	US-08-943-087-49	Sequence 49, Appl
	134	16	0.5	663	2	US-08-943-087-51	Sequence 51, Appl
	135	16	0.5	663	2	US-08-943-087-53	Sequence 53, Appl
	136	16	0.5	663	2	US-08-943-087-55	Sequence 55, Appl
	137	16	0.5	663	2	US-08-943-087-59	Sequence 59, Appl
	138	16	0.5	668	4	US-09-385-982-157	Sequence 157, App
	139	16	0.5	700	4	US-09-236-097-7	Sequence 7, Appli
c	140	16	0.5	704	3	US-09-009-156-1	Sequence 1, Appli
c	141	16	0.5	704	4	US-09-372-154-1	Sequence 1, Appli
	142	16	0.5	718	4	US-09-347-803-17	Sequence 17, Appl
	143	16	0.5	745	1	US-08-036-555B-163	Sequence 163, App
	144	16	0.5	745	1	US-08-469-569-163	Sequence 163, App
	145	16	0.5	745	1	US-08-249-322A-163	Sequence 163, App
	146	16	0.5	745	1	US-08-469-526A-163	Sequence 163, App
	147	16	0.5	745	2	US-08-734-591A-163	Sequence 163, App
	148	16	0.5	745	2	US-08-469-660-163	Sequence 163, App
	149	16	0.5	745	3	US-08-341-018-51	Sequence 51, Appl
	150	16	0.5	745	3	US-08-470-335-163	Sequence 163, App
	151	16	0.5	745	4	US-08-735-021-163	Sequence 163, App
	152	16	0.5	745	4	US-08-734-664A-163	Sequence 163, App
	153	16	0.5	745	4	US-08-470-339-163	Sequence 163, App
	154	16	0.5	745	4	US-08-467-602-163	Sequence 163, App
	155	16	0.5	745	5	PCT-US94-05083C-159	Sequence 159, App
	156	16	0.5	745	5	PCT-US95-06846A-163	Sequence 163, App
c	157	16	0.5	780	4	US-09-134-001C-1631	Sequence 1631, Ap
	158	16	0.5	788	3	US-08-822-264-2	Sequence 2, Appli
c	159	16	0.5	874	4	US-08-881-450A-24	Sequence 24, Appl
c	160	16	0.5	887	4	US-09-329-350-36	Sequence 36, Appl
c	161	16	0.5	1044	1	US-07-975-526-3	Sequence 3, Appli
c	162	16	0.5	1044	4	US-07-974-409C-425	Sequence 425, App
	163	16	0.5	1059	4	US-09-576-160B-9	Sequence 9, Appli
	164	16	0.5	1062	4	US-09-576-160B-12	Sequence 12, Appl
c	165	16	0.5	1104	4	US-09-134-001C-1331	Sequence 1331, Ap
	166	16	0.5	1197	4	US-09-491-362-6	Sequence 6, Appli
	167	16	0.5	1197	4	US-09-874-562-6	Sequence 6, Appli
	168	16	0.5	1209	4	US-08-791-115B-24	Sequence 24, Appl
	169	16	0.5	1222	4	US-09-333-611-13	Sequence 13, Appl
c	170	16	0.5	1233	1	US-08-254-922-1	Sequence 1, Appli
c	171	16	0.5	1233	1	US-08-286-748B-1	Sequence 1, Appli
c	172	16	0.5	1236	1	US-07-957-039A-7	Sequence 7, Appli
c	173	16	0.5	1236	1	US-08-153-799-17	Sequence 17, Appl
	174	16	0.5	1247	1	US-08-278-729A-32	Sequence 32, Appl
	175	16	0.5	1247	1	US-08-155-343A-32	Sequence 32, Appl
	176	16	0.5	1247	1	US-08-406-672-32	Sequence 32, Appl
	177	16	0.5	1247	1	US-08-643-563A-32	Sequence 32, Appl
	178	16	0.5	1247	1	US-08-643-763A-32	Sequence 32, Appl
	179	16	0.5	1247	1	US-08-462-623-32	Sequence 32, Appl
	180	16	0.5	1247	1	US-08-451-953A-32	Sequence 32, Appl
	181	16	0.5	1247	2	US-08-445-468A-32	Sequence 32, Appl
	182	16	0.5	1247	2	US-08-461-397A-32	Sequence 32, Appl
	183	16	0.5	1247	2	US-08-912-088-32	Sequence 32, Appl
	184	16	0.5	1247	3	US-08-278-730A-32	Sequence 32, Appl
	185	16	0.5	1247	3	US-08-445-467-32	Sequence 32, Appl
	186	16	0.5	1247	3	US-08-480-515A-32	Sequence 32, Appl
	187	16	0.5	1247	4	US-09-170-936-32	Sequence 32, Appl
	188	16	0.5	1247	4	US-08-461-113-32	Sequence 32, Appl
	189	16	0.5	1247	5	PCT-US93-07190-32	Sequence 32, Appl
	190	16	0.5	1247	5	PCT-US93-07231-32	Sequence 32, Appl
	191	16	0.5	1247	5	PCT-US93-08742-32	Sequence 32, Appl
	192	16	0.5	1247	5	PCT-US93-08808-32	Sequence 32, Appl
	193	16	0.5	1247	5	PCT-US93-08885-32	Sequence 32, Appl
	194	16	0.5	1254	4	US-09-322-478-25	Sequence 25, Appl
	195	16	0.5	1263	1	US-08-532-828B-11	Sequence 11, Appl
	196	16	0.5	1263	1	US-08-532-828B-12	Sequence 12, Appl
c	197	16	0.5	1294	4	US-09-312-038-4	Sequence 4, Appli
c	198	16	0.5	1335	4	US-09-859-698-1	Sequence 1, Appli
c	199	16	0.5	1335	4	US-09-428-979-1	Sequence 1, Appli
	200	16	0.5	1345	4	US-09-372-339-1	Sequence 1, Appli
	201	16	0.5	1345	4	US-09-372-339-2	Sequence 2, Appli
	202	16	0.5	1345	4	US-09-144-367-3	Sequence 3, Appli
c	203	16	0.5	1372	6	5219569-1	Patent No. 5219569
	204	16	0.5	1393	2	US-08-765-875-1	Sequence 1, Appli
	205	16	0.5	1393	3	US-08-795-671-1	Sequence 1, Appli
c	206	16	0.5	1401	4	US-09-134-001C-595	Sequence 595, App
	207	16	0.5	1402	2	US-09-205-860-1	Sequence 1, Appli
	208	16	0.5	1425	1	US-08-464-148-1	Sequence 1, Appli
	209	16	0.5	1425	1	US-08-385-500-1	Sequence 1, Appli
	210	16	0.5	1425	1	US-08-846-784-1	Sequence 1, Appli
	211	16	0.5	1434	4	US-09-480-921B-9	Sequence 9, Appli
	212	16	0.5	1434	4	US-09-480-921B-28	Sequence 28, Appl
c	213	16	0.5	1475	4	US-09-643-597-122	Sequence 122, App
	214	16	0.5	1477	3	US-08-096-181A-7	Sequence 7, Appli
	215	16	0.5	1477	5	PCT-US94-08326-7	Sequence 7, Appli
	216	16	0.5	1614	2	US-08-867-941-10	Sequence 10, Appl
	217	16	0.5	1614	4	US-09-074-658-10	Sequence 10, Appl
	218	16	0.5	1623	2	US-08-867-941-5	Sequence 5, Appli
	219	16	0.5	1623	4	US-09-074-658-5	Sequence 5, Appli
	220	16	0.5	1643	1	US-08-532-828B-1	Sequence 1, Appli
	221	16	0.5	1643	1	US-08-532-828B-2	Sequence 2, Appli
	222	16	0.5	1643	1	US-08-532-828B-7	Sequence 7, Appli
	223	16	0.5	1643	1	US-08-532-828B-8	Sequence 8, Appli
	224	16	0.5	1643	1	US-08-532-828B-9	Sequence 9, Appli
	225	16	0.5	1643	1	US-08-532-828B-10	Sequence 10, Appl
	226	16	0.5	1643	1	US-08-700-359-7	Sequence 7, Appli
	227	16	0.5	1643	1	US-08-700-359-8	Sequence 8, Appli
	228	16	0.5	1643	1	US-08-700-359-10	Sequence 10, Appl
	229	16	0.5	1643	1	US-08-674-168-21	Sequence 21, Appl
	230	16	0.5	1643	2	US-08-596-366-5	Sequence 5, Appli
	231	16	0.5	1643	2	US-08-596-366-7	Sequence 7, Appli
	232	16	0.5	1643	2	US-08-967-104-5	Sequence 5, Appli
	233	16	0.5	1643	2	US-08-967-104-7	Sequence 7, Appli
	234	16	0.5	1643	3	US-08-985-908-3	Sequence 3, Appli
	235	16	0.5	1643	3	US-08-985-908-4	Sequence 4, Appli
	236	16	0.5	1643	3	US-08-985-908-6	Sequence 6, Appli
	237	16	0.5	1643	3	US-08-852-730-12	Sequence 12, Appl
	238	16	0.5	1643	3	US-08-852-730-13	Sequence 13, Appl
	239	16	0.5	1643	3	US-08-852-730-15	Sequence 15, Appl
	240	16	0.5	1643	4	US-08-985-916-3	Sequence 3, Appli
	241	16	0.5	1643	4	US-08-985-916-4	Sequence 4, Appli
	242	16	0.5	1643	4	US-08-985-916-6	Sequence 6, Appli
	243	16	0.5	1659	2	US-08-943-087-13	Sequence 13, Appl
	244	16	0.5	1659	2	US-08-943-087-15	Sequence 15, Appl
	245	16	0.5	1659	2	US-08-943-087-25	Sequence 25, Appl
	246	16	0.5	1659	2	US-08-943-087-27	Sequence 27, Appl

	247	16	0.5	1659	2	US-08-943-087-29	Sequence 29, Appl
	248	16	0.5	1659	2	US-08-943-087-31	Sequence 31, Appl
	249	16	0.5	1659	2	US-08-943-087-33	Sequence 33, Appl
	250	16	0.5	1659	2	US-08-943-087-35	Sequence 35, Appl
	251	16	0.5	1659	2	US-08-943-087-37	Sequence 37, Appl
	252	16	0.5	1659	2	US-08-943-087-39	Sequence 39, Appl
	253	16	0.5	1659	2	US-08-943-087-41	Sequence 41, Appl
	254	16	0.5	1659	2	US-08-943-087-43	Sequence 43, Appl
	255	16	0.5	1659	2	US-08-943-087-45	Sequence 45, Appl
c	256	16	0.5	1674	2	US-08-999-811-1	Sequence 1, Appli
c	257	16	0.5	1674	3	US-09-042-105-1	Sequence 1, Appli
c	258	16	0.5	1674	4	US-09-134-001C-1555	Sequence 1555, Ap
c	259	16	0.5	1674	5	PCT-US96-09001-1	Sequence 1, Appli
c	260	16	0.5	1702	1	US-07-616-022C-1	Sequence 1, Appli
	261	16	0.5	1732	4	US-09-449-335-1	Sequence 1, Appli
	262	16	0.5	1732	4	US-09-449-335-5	Sequence 5, Appli
c	263	16	0.5	1809	1	US-08-455-001-1	Sequence 1, Appli
c	264	16	0.5	1809	5	PCT-US95-11869-1	Sequence 1, Appli
	265	16	0.5	1837	2	US-08-750-134A-4	Sequence 4, Appli
	266	16	0.5	1837	4	US-09-363-745-4	Sequence 4, Appli
	267	16	0.5	1868	2	US-08-960-022-5	Sequence 5, Appli
c	268	16	0.5	1882	1	US-08-696-349-1	Sequence 1, Appli
c	269	16	0.5	1882	5	PCT-US96-13156-1	Sequence 1, Appli
	270	16	0.5	1886	1	US-07-980-526-1	Sequence 1, Appli
	271	16	0.5	1894	1	US-08-476-008-66	Sequence 66, Appl
	272	16	0.5	1894	1	US-08-306-063-66	Sequence 66, Appl
	273	16	0.5	1894	1	US-08-833-485-66	Sequence 66, Appl
	274	16	0.5	1894	4	US-09-137-440-66	Sequence 66, Appl
c	275	16	0.5	1908	2	US-08-909-965C-17	Sequence 17, Appl
c	276	16	0.5	1929	4	US-09-040-229B-7	Sequence 7, Appli
c	277	16	0.5	1953	1	US-08-315-468-5	Sequence 5, Appli
c	278	16	0.5	1953	4	US-07-941-650A-3	Sequence 3, Appli
	279	16	0.5	1988	2	US-08-712-948-8	Sequence 8, Appli
c	280	16	0.5	1996	1	US-08-385-370-1	Sequence 1, Appli
c	281	16	0.5	1997	3	US-08-795-430-7	Sequence 7, Appli
c	282	16	0.5	1997	4	US-08-510-133A-34	Sequence 34, Appl
c	283	16	0.5	1997	4	US-09-355-700-7	Sequence 7, Appli
c	284	16	0.5	1997	4	US-08-601-132-32	Sequence 32, Appl
	285	16	0.5	2003	1	US-08-036-555B-21	Sequence 21, Appl
	286	16	0.5	2003	1	US-08-469-569-21	Sequence 21, Appl
	287	16	0.5	2003	1	US-08-249-322A-21	Sequence 21, Appl
	288	16	0.5	2003	1	US-08-469-526A-21	Sequence 21, Appl
	289	16	0.5	2003	2	US-08-734-591A-21	Sequence 21, Appl
	290	16	0.5	2003	2	US-08-469-660-21	Sequence 21, Appl
	291	16	0.5	2003	3	US-08-341-018-71	Sequence 71, Appl
	292	16	0.5	2003	3	US-08-470-335-21	Sequence 21, Appl
	293	16	0.5	2003	4	US-08-735-021-21	Sequence 21, Appl
	294	16	0.5	2003	4	US-08-734-664A-21	Sequence 21, Appl
	295	16	0.5	2003	4	US-08-470-339-21	Sequence 21, Appl
	296	16	0.5	2003	4	US-08-467-602-21	Sequence 21, Appl
	297	16	0.5	2003	5	PCT-US94-05083C-21	Sequence 21, Appl
	298	16	0.5	2003	5	PCT-US95-06846A-21	Sequence 21, Appl
	299	16	0.5	2007	4	US-09-149-476-135	Sequence 135, App
	300	16	0.5	2009	4	US-09-446-504-17	Sequence 17, Appl
	301	16	0.5	2009	4	US-09-712-266-17	Sequence 17, Appl
c	302	16	0.5	2023	4	US-09-491-522-6	Sequence 6, Appli
c	303	16	0.5	2031	4	US-08-706-054A-1	Sequence 1, Appli
	304	16	0.5	2031	4	US-08-706-054A-2	Sequence 2, Appli
c	305	16	0.5	2077	1	US-08-217-327-7	Sequence 7, Appli
	306	16	0.5	2117	1	US-08-431-080-23	Sequence 23, Appl
	307	16	0.5	2117	2	US-08-938-534-23	Sequence 23, Appl
	308	16	0.5	2117	4	US-09-345-294-23	Sequence 23, Appl
	309	16	0.5	2156	2	US-08-959-011-2	Sequence 2, Appli
	310	16	0.5	2180	2	US-08-755-559-2	Sequence 2, Appli
	311	16	0.5	2180	3	US-09-210-474-2	Sequence 2, Appli
	312	16	0.5	2180	4	US-09-539-774-2	Sequence 2, Appli
c	313	16	0.5	2190	4	US-08-893-654B-1	Sequence 1, Appli
c	314	16	0.5	2263	1	US-08-176-126B-1	Sequence 1, Appli
c	315	16	0.5	2263	2	US-08-669-435-1	Sequence 1, Appli
c	316	16	0.5	2263	5	PCT-US94-14431A-1	Sequence 1, Appli
	317	16	0.5	2268	4	US-08-675-773B-4	Sequence 4, Appli
c	318	16	0.5	2289	4	US-09-312-038-3	Sequence 3, Appli
c	319	16	0.5	2294	4	US-09-643-597-123	Sequence 123, App

c	320	16	0.5	2301	6	5188829-2	Patent No. 5188829
c	321	16	0.5	2323	4	US-09-149-476-24	Sequence 24, Appl
	322	16	0.5	2335	2	US-08-300-584-3	Sequence 3, Appli
	323	16	0.5	2335	4	US-08-476-123-3	Sequence 3, Appli
	324	16	0.5	2360	1	US-08-039-364-1	Sequence 1, Appli
	325	16	0.5	2360	4	US-09-158-710-1	Sequence 1, Appli
c	326	16	0.5	2450	4	US-09-491-522-2	Sequence 2, Appli
c	327	16	0.5	2456	1	US-07-882-711-1	Sequence 1, Appli
c	328	16	0.5	2456	2	US-08-462-174-1	Sequence 1, Appli
c	329	16	0.5	2457	2	US-08-723-415B-12	Sequence 12, Appl
c	330	16	0.5	2457	3	US-09-189-627A-12	Sequence 12, Appl
c	331	16	0.5	2457	4	US-09-710-861-12	Sequence 12, Appl
c	332	16	0.5	2492	1	US-08-139-937-13	Sequence 13, Appl
c	333	16	0.5	2492	5	PCT-US93-11310-13	Sequence 13, Appl
c	334	16	0.5	2517	1	US-08-306-691B-18	Sequence 18, Appl
c	335	16	0.5	2517	1	US-08-385-142-2	Sequence 2, Appli
c	336	16	0.5	2517	2	US-08-481-814A-1	Sequence 1, Appli
c	337	16	0.5	2517	3	US-08-801-092-2	Sequence 2, Appli
c	338	16	0.5	2517	4	US-09-517-584A-3	Sequence 3, Appli
c	339	16	0.5	2517	4	US-09-315-113-2	Sequence 2, Appli
c	340	16	0.5	2572	4	US-09-221-017B-225	Sequence 225, App
	341	16	0.5	2671	2	US-08-408-519-1	Sequence 1, Appli
	342	16	0.5	2671	5	PCT-US95-03552-1	Sequence 1, Appli
	343	16	0.5	2730	2	US-08-811-897A-39	Sequence 39, Appl
	344	16	0.5	2730	2	US-08-855-213-39	Sequence 39, Appl
	345	16	0.5	2730	4	US-09-201-474-39	Sequence 39, Appl
c	346	16	0.5	2745	1	US-08-385-370-3	Sequence 3, Appli
	347	16	0.5	2814	2	US-08-811-897A-38	Sequence 38, Appl
	348	16	0.5	2814	2	US-08-855-213-38	Sequence 38, Appl
	349	16	0.5	2814	4	US-09-201-474-38	Sequence 38, Appl
c	350	16	0.5	2814	4	US-09-177-650-90	Sequence 90, Appl
c	351	16	0.5	2832	3	US-08-819-177-2	Sequence 2, Appli
c	352	16	0.5	2838	4	US-08-246-489-1	Sequence 1, Appli
	353	16	0.5	2850	2	US-08-224-482-7	Sequence 7, Appli
c	354	16	0.5	2935	4	US-09-480-921B-27	Sequence 27, Appl
c	355	16	0.5	2953	2	US-08-859-201-1	Sequence 1, Appli
c	356	16	0.5	2965	3	US-08-906-360-2	Sequence 2, Appli
	357	16	0.5	3123	1	US-08-094-889-2	Sequence 2, Appli
c	358	16	0.5	3200	1	US-08-444-405-1	Sequence 1, Appli
c	359	16	0.5	3200	1	US-08-384-850-1	Sequence 1, Appli
	360	16	0.5	3278	1	US-08-484-105-13	Sequence 13, Appl
	361	16	0.5	3278	1	US-08-484-106-13	Sequence 13, Appl
c	362	16	0.5	3358	2	US-08-469-537A-104	Sequence 104, App
	363	16	0.5	3516	2	US-08-943-087-1	Sequence 1, Appli
	364	16	0.5	3564	4	US-09-347-878-15	Sequence 15, Appl
c	365	16	0.5	3566	2	US-08-415-788-1	Sequence 1, Appli
c	366	16	0.5	3566	2	US-08-415-788-4	Sequence 4, Appli
c	367	16	0.5	3566	2	US-08-415-788-6	Sequence 6, Appli
c	368	16	0.5	3567	2	US-08-415-788-2	Sequence 2, Appli
c	369	16	0.5	3567	4	US-08-637-732A-3	Sequence 3, Appli
c	370	16	0.5	3627	1	US-08-104-072B-6	Sequence 6, Appli
c	371	16	0.5	3627	1	US-08-351-413-7	Sequence 7, Appli
c	372	16	0.5	3627	2	US-09-025-583-7	Sequence 7, Appli
	373	16	0.5	3639	2	US-08-737-524B-26	Sequence 26, Appl
c	374	16	0.5	3639	2	US-08-737-524B-26	Sequence 26, Appl
	375	16	0.5	3651	4	US-07-852-132A-13	Sequence 13, Appl
	376	16	0.5	3666	6	5248670-3	Patent No. 5248670
	377	16	0.5	3765	3	US-07-705-490-1	Sequence 1, Appli
	378	16	0.5	3765	4	US-07-751-891B-1	Sequence 1, Appli
c	379	16	0.5	3989	1	US-08-327-494A-1	Sequence 1, Appli
c	380	16	0.5	3989	1	US-08-327-494A-3	Sequence 3, Appli
c	381	16	0.5	3989	5	PCT-US95-13659-1	Sequence 1, Appli
c	382	16	0.5	3989	5	PCT-US95-13659-3	Sequence 3, Appli
	383	16	0.5	4010	2	US-08-785-310A-3	Sequence 3, Appli
	384	16	0.5	4098	4	US-09-268-866-1	Sequence 1, Appli
	385	16	0.5	4134	4	US-09-162-021B-1	Sequence 1, Appli
	386	16	0.5	4134	4	US-09-687-477-17	Sequence 17, Appl
	387	16	0.5	4362	2	US-08-455-073A-1	Sequence 1, Appli
	388	16	0.5	4402	4	US-09-484-970B-135	Sequence 135, App
c	389	16	0.5	4804	4	US-09-066-047-6	Sequence 6, Appli
	390	16	0.5	4935	2	US-08-631-097-3	Sequence 3, Appli
	391	16	0.5	5288	2	US-08-540-406-18	Sequence 18, Appl
	392	16	0.5	5288	3	US-08-656-055-18	Sequence 18, Appl

	393	16	0.5	5288	4	US-08-954-668-18	Sequence 18, Appl
	394	16	0.5	5288	4	US-08-918-658-18	Sequence 18, Appl
	395	16	0.5	5288	5	PCT-US95-13233-18	Sequence 18, Appl
	396	16	0.5	5886	4	US-08-810-712-9	Sequence 9, Appli
	397	16	0.5	6211	4	US-08-961-527-8	Sequence 8, Appli
c	398	16	0.5	6370	4	US-09-306-595C-2	Sequence 2, Appli
c	399	16	0.5	6692	4	US-09-491-522-1	Sequence 1, Appli
	400	16	0.5	6714	4	US-09-299-141-6	Sequence 6, Appli
	401	16	0.5	6803	3	US-08-665-259-19	Sequence 19, Appl
	402	16	0.5	6803	3	US-08-762-500-19	Sequence 19, Appl
	403	16	0.5	6924	4	US-09-299-141-9	Sequence 9, Appli
	404	16	0.5	6924	4	US-09-299-141-10	Sequence 10, Appl
	405	16	0.5	6924	4	US-09-299-141-11	Sequence 11, Appl
c	406	16	0.5	7176	4	US-09-221-017B-626	Sequence 626, App
	407	16	0.5	7336	4	US-09-306-446C-1	Sequence 1, Appli
	408	16	0.5	7641	2	US-08-867-941-6	Sequence 6, Appli
	409	16	0.5	7641	4	US-09-074-658-6	Sequence 6, Appli
	410	16	0.5	7650	2	US-08-867-941-1	Sequence 1, Appli
	411	16	0.5	7650	4	US-09-074-658-1	Sequence 1, Appli
c	412	16	0.5	8040	1	US-08-596-291-1	Sequence 1, Appli
c	413	16	0.5	8040	3	US-09-100-804-1	Sequence 1, Appli
c	414	16	0.5	8043	5	PCT-US94-09943-1	Sequence 1, Appli
c	415	16	0.5	8050	4	US-09-491-362-11	Sequence 11, Appl
c	416	16	0.5	8050	4	US-09-874-562-11	Sequence 11, Appl
c	417	16	0.5	8119	4	US-09-290-640-45	Sequence 45, Appl
	418	16	0.5	8906	2	US-08-826-267-1	Sequence 1, Appli
	419	16	0.5	8982	3	US-08-976-255-5	Sequence 5, Appli
	420	16	0.5	9837	4	US-09-221-017B-636	Sequence 636, App
c	421	16	0.5	9972	3	US-08-836-022A-3	Sequence 3, Appli
c	422	16	0.5	9972	4	US-09-427-048A-3	Sequence 3, Appli
	423	16	0.5	10299	2	US-08-477-451-1	Sequence 1, Appli
c	424	16	0.5	10299	2	US-08-477-451-5	Sequence 5, Appli
c	425	16	0.5	10348	2	US-08-457-273B-41	Sequence 41, Appl
c	426	16	0.5	10348	3	US-08-556-419-13	Sequence 13, Appl
c	427	16	0.5	10348	4	US-09-041-886-14	Sequence 14, Appl
c	428	16	0.5	10366	1	US-08-246-982A-5	Sequence 5, Appli
c	429	16	0.5	10366	1	US-08-453-265-5	Sequence 5, Appli
	430	16	0.5	11580	4	US-09-334-220-4	Sequence 4, Appli
	431	16	0.5	19932	2	US-08-477-451-25	Sequence 25, Appl
	432	16	0.5	35408	4	US-08-973-334-3	Sequence 3, Appli
	433	16	0.5	35408	4	US-09-563-869A-3	Sequence 3, Appli
	434	16	0.5	35408	4	US-08-549-489-3	Sequence 3, Appli
	435	16	0.5	35524	3	US-08-923-137-1	Sequence 1, Appli
	436	16	0.5	72604	4	US-09-268-992-7	Sequence 7, Appli
	437	16	0.5	72604	4	US-09-657-474-7	Sequence 7, Appli
	438	16	0.5	87350	3	US-08-781-891-79	Sequence 79, Appl
	439	16	0.5	87543	4	US-09-791-211-3	Sequence 3, Appli
	440	15	0.5	15	2	US-08-863-639A-21	Sequence 21, Appl
c	441	15	0.5	17	3	US-08-909-742-3	Sequence 3, Appli
c	442	15	0.5	17	3	US-08-909-742-4	Sequence 4, Appli
c	443	15	0.5	17	4	US-09-412-289-3	Sequence 3, Appli
c	444	15	0.5	17	4	US-09-412-289-4	Sequence 4, Appli
	445	15	0.5	18	2	US-08-857-946-14	Sequence 14, Appl
	446	15	0.5	18	3	US-08-970-740-14	Sequence 14, Appl
	447	15	0.5	18	3	US-09-143-212-45	Sequence 45, Appl
	448	15	0.5	20	2	US-08-465-485A-28	Sequence 28, Appl
	449	15	0.5	20	3	US-09-080-285-28	Sequence 28, Appl
	450	15	0.5	20	4	US-09-030-701-65	Sequence 65, Appl
	451	15	0.5	20	4	US-09-593-711A-37	Sequence 37, Appl
	452	15	0.5	20	4	US-09-082-649B-57	Sequence 57, Appl
	453	15	0.5	20	4	US-09-724-426-28	Sequence 28, Appl
c	454	15	0.5	21	2	US-08-863-639A-52	Sequence 52, Appl
c	455	15	0.5	21	2	US-08-863-639A-55	Sequence 55, Appl
	456	15	0.5	21	2	US-08-863-639A-56	Sequence 56, Appl
c	457	15	0.5	21	2	US-08-863-639A-67	Sequence 67, Appl
	458	15	0.5	21	2	US-08-863-639A-68	Sequence 68, Appl
	459	15	0.5	21	2	US-08-863-639A-71	Sequence 71, Appl
	460	15	0.5	21	2	US-08-416-214A-11	Sequence 11, Appl
c	461	15	0.5	24	2	US-08-570-155-16	Sequence 16, Appl
	462	15	0.5	24	2	US-08-570-155-17	Sequence 17, Appl
	463	15	0.5	25	1	US-08-374-144-3	Sequence 3, Appli
	464	15	0.5	25	1	US-08-775-164-3	Sequence 3, Appli
	465	15	0.5	25	2	US-08-775-609-3	Sequence 3, Appli
	466	15	0.5	25	2	US-08-775-607-3	Sequence 3, Appli
	467	15	0.5	25	3	US-08-622-277A-17	Sequence 17, Appl
	468	15	0.5	25	5	PCT-US93-06828-3	Sequence 3, Appli
	469	15	0.5	28	3	US-08-863-813A-38	Sequence 38, Appl
	470	15	0.5	28	4	US-09-025-580-6	Sequence 6, Appli
c	471	15	0.5	28	4	US-08-676-318A-38	Sequence 38, Appl
	472	15	0.5	30	1	US-08-246-978A-2	Sequence 2, Appli
c	473	15	0.5	30	1	US-08-068-747-2	Sequence 2, Appli
	474	15	0.5	30	1	US-08-440-814A-2	Sequence 2, Appli
	475	15	0.5	33	1	US-08-068-747-7	Sequence 7, Appli
	476	15	0.5	38	3	US-08-863-813A-51	Sequence 51, Appl
c	477	15	0.5	41	4	US-09-402-631A-1	Sequence 1, Appli
c	478	15	0.5	45	4	US-09-497-933A-22	Sequence 22, Appl
	479	15	0.5	48	4	US-09-497-933A-19	Sequence 19, Appl
	480	15	0.5	50	4	US-09-930-181-13	Sequence 13, Appl
	481	15	0.5	50	4	US-09-930-181-14	Sequence 14, Appl
	482	15	0.5	60	4	US-09-497-933A-21	Sequence 21, Appl
c	483	15	0.5	63	4	US-09-497-933A-18	Sequence 18, Appl
	484	15	0.5	63	4	US-09-497-933A-23	Sequence 23, Appl
	485	15	0.5	63	4	US-09-497-933A-25	Sequence 25, Appl
	486	15	0.5	65	3	US-08-795-430-17	Sequence 17, Appl
	487	15	0.5	65	4	US-09-355-700-17	Sequence 17, Appl
	488	15	0.5	66	3	US-08-516-859A-96	Sequence 96, Appl
	489	15	0.5	66	4	US-09-586-472-96	Sequence 96, Appl
	490	15	0.5	66	4	US-09-528-706-96	Sequence 96, Appl
c	491	15	0.5	69	3	US-08-480-173A-31	Sequence 31, Appl
	492	15	0.5	69	3	US-08-480-173A-32	Sequence 32, Appl
c	493	15	0.5	69	3	US-08-484-408A-31	Sequence 31, Appl
	494	15	0.5	69	3	US-08-484-408A-32	Sequence 32, Appl
c	495	15	0.5	77	2	US-08-180-524-5	Sequence 5, Appli
c	496	15	0.5	77	2	US-08-975-166-5	Sequence 5, Appli
c	497	15	0.5	78	4	US-09-497-933A-24	Sequence 24, Appl
	498	15	0.5	81	4	US-09-497-933A-20	Sequence 20, Appl
	499	15	0.5	84	3	US-08-795-430-19	Sequence 19, Appl
	500	15	0.5	84	4	US-09-355-700-19	Sequence 19, Appl
	501	15	0.5	90	1	US-08-398-617-15	Sequence 15, Appl
	502	15	0.5	90	2	US-08-398-615-15	Sequence 15, Appl
	503	15	0.5	90	4	US-07-827-691A-4	Sequence 4, Appli
	504	15	0.5	90	4	US-08-397-303-15	Sequence 15, Appl
	505	15	0.5	96	4	US-09-025-769B-291	Sequence 291, App
c	506	15	0.5	103	3	US-08-480-173A-12	Sequence 12, Appl
c	507	15	0.5	103	3	US-08-484-408A-12	Sequence 12, Appl
c	508	15	0.5	146	2	US-08-180-524-3	Sequence 3, Appli
c	509	15	0.5	146	2	US-08-975-166-3	Sequence 3, Appli
c	510	15	0.5	161	4	US-09-404-879A-35	Sequence 35, Appl
c	511	15	0.5	168	4	US-09-328-111-222	Sequence 222, App
c	512	15	0.5	180	1	US-08-343-281A-20	Sequence 20, Appl
	513	15	0.5	186	3	US-09-188-930-95	Sequence 95, Appl
	514	15	0.5	186	4	US-09-402-328-4	Sequence 4, Appli
	515	15	0.5	189	4	US-09-568-816A-3	Sequence 3, Appli
	516	15	0.5	192	4	US-09-134-001C-49	Sequence 49, Appl
	517	15	0.5	195	1	US-08-324-243-20	Sequence 20, Appl
	518	15	0.5	195	1	US-08-532-390-20	Sequence 20, Appl
	519	15	0.5	195	3	US-08-717-294-20	Sequence 20, Appl
	520	15	0.5	195	5	PCT-US95-11511-20	Sequence 20, Appl
c	521	15	0.5	200	4	US-08-983-502-3	Sequence 3, Appli
c	522	15	0.5	200	5	PCT-US96-10521-3	Sequence 3, Appli
c	523	15	0.5	249	1	US-07-872-644-40	Sequence 40, Appl
c	524	15	0.5	249	1	US-08-297-494-40	Sequence 40, Appl
c	525	15	0.5	249	1	US-08-297-510-40	Sequence 40, Appl
c	526	15	0.5	249	1	US-08-479-532-40	Sequence 40, Appl
c	527	15	0.5	249	1	US-08-455-526-40	Sequence 40, Appl
c	528	15	0.5	249	1	US-08-455-525-40	Sequence 40, Appl
c	529	15	0.5	249	3	US-09-139-491-40	Sequence 40, Appl
c	530	15	0.5	249	5	PCT-US92-03222-40	Sequence 40, Appl
c	531	15	0.5	256	4	US-08-483-533-11	Sequence 11, Appl
c	532	15	0.5	256	4	US-09-283-471A-11	Sequence 11, Appl
	533	15	0.5	261	4	US-09-009-816-7	Sequence 7, Appli
	534	15	0.5	288	2	US-08-716-942-16	Sequence 16, Appl
	535	15	0.5	288	4	US-09-130-337A-16	Sequence 16, Appl
	536	15	0.5	318	2	US-08-646-981-4	Sequence 4, Appli
	537	15	0.5	329	1	US-08-148-910-2	Sequence 2, Appli
	538	15	0.5	329	1	US-08-148-910-13	Sequence 13, Appl

	539	15	0.5	329	1	US-08-448-937A-2	Sequence 2, Appli
	540	15	0.5	329	1	US-08-448-937A-13	Sequence 13, Appl
c	541	15	0.5	330	4	US-09-325-932A-93	Sequence 93, Appl
	542	15	0.5	346	2	US-08-761-277A-50	Sequence 50, Appl
	543	15	0.5	349	4	US-09-198-119C-94	Sequence 94, Appl
	544	15	0.5	364	4	US-09-221-017B-481	Sequence 481, App
c	545	15	0.5	367	4	US-09-325-932A-2	Sequence 2, Appli
c	546	15	0.5	423	1	US-08-470-179-190	Sequence 190, App
c	547	15	0.5	425	1	US-08-664-596B-12	Sequence 12, Appl
c	548	15	0.5	425	4	US-09-397-787-249	Sequence 249, App
c	549	15	0.5	431	4	US-08-483-533-17	Sequence 17, Appl
c	550	15	0.5	431	4	US-09-283-471A-17	Sequence 17, Appl
	551	15	0.5	450	4	US-09-370-838-145	Sequence 145, App
c	552	15	0.5	457	3	US-08-952-664-10	Sequence 10, Appl
c	553	15	0.5	457	4	US-09-487-874-10	Sequence 10, Appl
c	554	15	0.5	463	4	US-09-221-017B-201	Sequence 201, App
c	555	15	0.5	465	4	US-09-171-209-29	Sequence 29, Appl
	556	15	0.5	500	4	US-09-575-574-2	Sequence 2, Appli
c	557	15	0.5	501	4	US-09-404-879A-118	Sequence 118, App
c	558	15	0.5	502	4	US-08-998-416-124	Sequence 124, App
	559	15	0.5	506	1	US-08-398-617-13	Sequence 13, Appl
	560	15	0.5	506	2	US-08-398-615-13	Sequence 13, Appl
	561	15	0.5	506	4	US-08-397-303-13	Sequence 13, Appl
	562	15	0.5	510	4	US-09-452-239-9	Sequence 9, Appli
	563	15	0.5	516	2	US-08-943-814-1	Sequence 1, Appli
	564	15	0.5	519	2	US-08-581-528A-5	Sequence 5, Appli
	565	15	0.5	519	5	PCT-US94-07799-5	Sequence 5, Appli
	566	15	0.5	530	3	US-08-758-662-4	Sequence 4, Appli
	567	15	0.5	531	4	US-09-191-608-8	Sequence 8, Appli
	568	15	0.5	539	2	US-08-702-703-1	Sequence 1, Appli
	569	15	0.5	541	2	US-08-943-814-9	Sequence 9, Appli
	570	15	0.5	543	1	US-08-468-036-11	Sequence 11, Appl
	571	15	0.5	543	2	US-08-376-843-11	Sequence 11, Appl
	572	15	0.5	551	2	US-09-109-266-3	Sequence 3, Appli
	573	15	0.5	564	4	US-09-134-001C-56	Sequence 56, Appl
c	574	15	0.5	572	4	US-09-449-285A-12	Sequence 12, Appl
c	575	15	0.5	582	1	US-08-722-001-23	Sequence 23, Appl
c	576	15	0.5	587	4	US-09-221-017B-783	Sequence 783, App
	577	15	0.5	593	1	US-08-443-568B-13	Sequence 13, Appl
	578	15	0.5	593	5	PCT-US94-06997-13	Sequence 13, Appl
c	579	15	0.5	594	4	US-09-280-116-240	Sequence 240, App
	580	15	0.5	595	2	US-09-109-266-15	Sequence 15, Appl
	581	15	0.5	595	4	US-09-228-986-42	Sequence 42, Appl
	582	15	0.5	615	4	US-08-998-416-1063	Sequence 1063, Ap
c	583	15	0.5	633	1	US-08-234-783-1	Sequence 1, Appli
c	584	15	0.5	633	1	US-08-456-907-1	Sequence 1, Appli
	585	15	0.5	633	3	US-08-950-925-1	Sequence 1, Appli
c	586	15	0.5	633	5	PCT-US95-05523-1	Sequence 1, Appli
	587	15	0.5	635	4	US-09-385-982-3	Sequence 3, Appli
	588	15	0.5	640	2	US-08-835-099A-16	Sequence 16, Appl
	589	15	0.5	640	3	US-09-157-349-16	Sequence 16, Appl
	590	15	0.5	645	3	US-08-985-526-22	Sequence 22, Appl
	591	15	0.5	645	4	US-09-535-008-30	Sequence 30, Appl
	592	15	0.5	651	2	US-08-961-858-1	Sequence 1, Appli
	593	15	0.5	651	2	US-08-961-858-3	Sequence 3, Appli
	594	15	0.5	651	3	US-09-089-593-1	Sequence 1, Appli
	595	15	0.5	651	3	US-09-089-593-3	Sequence 3, Appli
	596	15	0.5	651	3	US-08-993-380-5	Sequence 5, Appli
	597	15	0.5	666	1	US-08-663-023-16	Sequence 16, Appl
c	598	15	0.5	670	4	US-09-009-816-3	Sequence 3, Appli
c	599	15	0.5	674	3	US-09-069-811-1	Sequence 1, Appli
c	600	15	0.5	674	4	US-09-543-106-1	Sequence 1, Appli
	601	15	0.5	711	4	US-08-998-416-1143	Sequence 1143, Ap
c	602	15	0.5	737	1	US-08-230-574-3	Sequence 3, Appli
	603	15	0.5	769	4	US-09-535-008-50	Sequence 50, Appl
	604	15	0.5	789	4	US-09-020-956-32	Sequence 32, Appl
	605	15	0.5	789	4	US-09-030-607-32	Sequence 32, Appl
	606	15	0.5	789	4	US-09-605-785-32	Sequence 32, Appl
	607	15	0.5	789	4	US-09-439-313-32	Sequence 32, Appl
	608	15	0.5	789	4	US-09-352-616A-32	Sequence 32, Appl
	609	15	0.5	789	4	US-09-232-149A-32	Sequence 32, Appl
	610	15	0.5	794	2	US-08-675-921B-8	Sequence 8, Appli
c	611	15	0.5	800	1	US-08-230-574-1	Sequence 1, Appli
	612	15	0.5	800	3	US-09-027-449-69	Sequence 69, Appl
	613	15	0.5	800	3	US-09-026-985-69	Sequence 69, Appl
	614	15	0.5	800	4	US-09-121-952A-69	Sequence 69, Appl
	615	15	0.5	800	4	US-09-234-340A-69	Sequence 69, Appl
	616	15	0.5	804	2	US-08-835-099A-10	Sequence 10, Appl
	617	15	0.5	804	3	US-09-157-349-10	Sequence 10, Appl
c	618	15	0.5	819	3	US-08-792-014-2	Sequence 2, Appli
c	619	15	0.5	819	4	US-09-443-948-2	Sequence 2, Appli
	620	15	0.5	826	4	US-08-998-416-364	Sequence 364, App
	621	15	0.5	829	3	US-08-961-083-133	Sequence 133, App
	622	15	0.5	832	2	US-08-686-417-1	Sequence 1, Appli
	623	15	0.5	834	4	US-09-134-001C-133	Sequence 133, App
	624	15	0.5	836	3	US-08-674-984-1	Sequence 1, Appli
	625	15	0.5	836	3	US-08-674-984-2	Sequence 2, Appli
	626	15	0.5	836	5	PCT-US95-15601-1	Sequence 1, Appli
	627	15	0.5	836	5	PCT-US95-15601-2	Sequence 2, Appli
c	628	15	0.5	856	2	US-09-057-762-23	Sequence 23, Appl
c	629	15	0.5	856	3	US-08-326-119A-23	Sequence 23, Appl
c	630	15	0.5	859	4	US-09-177-650-117	Sequence 117, App
c	631	15	0.5	862	4	US-08-818-112-9	Sequence 9, Appli
c	632	15	0.5	862	4	US-08-818-111-9	Sequence 9, Appli
c	633	15	0.5	862	4	US-09-056-556-9	Sequence 9, Appli
c	634	15	0.5	862	4	US-09-072-596-9	Sequence 9, Appli
	635	15	0.5	869	1	US-08-249-671A-7	Sequence 7, Appli
	636	15	0.5	875	2	US-08-778-912A-5	Sequence 5, Appli
	637	15	0.5	875	4	US-09-541-941B-5	Sequence 5, Appli
	638	15	0.5	879	1	US-08-249-671A-10	Sequence 10, Appl
c	639	15	0.5	879	2	US-09-024-848-1	Sequence 1, Appli
c	640	15	0.5	879	4	US-09-348-116A-1	Sequence 1, Appli
c	641	15	0.5	885	3	US-08-545-196B-20	Sequence 20, Appl
	642	15	0.5	887	4	US-09-329-350-36	Sequence 36, Appl
	643	15	0.5	897	2	US-08-486-663A-19	Sequence 19, Appl
	644	15	0.5	897	3	US-08-767-942A-24	Sequence 24, Appl
	645	15	0.5	897	4	US-09-134-001C-645	Sequence 645, App
	646	15	0.5	903	4	US-09-134-001C-2489	Sequence 2489, Ap
c	647	15	0.5	906	4	US-08-957-351-6	Sequence 6, Appli
	648	15	0.5	906	4	US-09-453-702B-117	Sequence 117, App
	649	15	0.5	906	4	US-09-134-001C-78	Sequence 78, Appl
c	650	15	0.5	912	4	US-09-105-390-39	Sequence 39, Appl
	651	15	0.5	912	4	US-09-134-001C-340	Sequence 340, App
	652	15	0.5	919	4	US-09-342-084-9	Sequence 9, Appli
c	653	15	0.5	925	2	US-08-916-901-4	Sequence 4, Appli
c	654	15	0.5	925	4	US-09-154-602-4	Sequence 4, Appli
	655	15	0.5	925	4	US-09-257-179-17	Sequence 17, Appl
	656	15	0.5	927	3	US-09-027-449-59	Sequence 59, Appl
	657	15	0.5	927	3	US-08-804-444A-59	Sequence 59, Appl
	658	15	0.5	927	3	US-09-026-985-59	Sequence 59, Appl
	659	15	0.5	927	4	US-09-121-952A-59	Sequence 59, Appl
	660	15	0.5	927	4	US-09-234-340A-59	Sequence 59, Appl
c	661	15	0.5	930	4	US-09-105-058C-17	Sequence 17, Appl
	662	15	0.5	936	4	US-09-416-509C-2	Sequence 2, Appli
c	663	15	0.5	969	4	US-09-134-001C-1227	Sequence 1227, Ap
	664	15	0.5	970	1	US-08-148-910-3	Sequence 3, Appli
	665	15	0.5	970	1	US-08-448-937A-3	Sequence 3, Appli
	666	15	0.5	991	4	US-09-344-529-3	Sequence 3, Appli
c	667	15	0.5	993	4	US-09-105-390-55	Sequence 55, Appl
c	668	15	0.5	1001	4	US-09-641-638-159	Sequence 159, App
c	669	15	0.5	1017	4	US-08-957-351-5	Sequence 5, Appli
c	670	15	0.5	1023	1	US-08-252-966B-16	Sequence 16, Appl
	671	15	0.5	1026	4	US-07-751-891B-24	Sequence 24, Appl
	672	15	0.5	1027	4	US-09-344-529-1	Sequence 1, Appli
	673	15	0.5	1028	4	US-08-118-200-1	Sequence 1, Appli
	674	15	0.5	1028	4	US-08-458-745-1	Sequence 1, Appli
	675	15	0.5	1037	1	US-08-240-124-1	Sequence 1, Appli
	676	15	0.5	1037	1	US-08-453-943-1	Sequence 1, Appli
	677	15	0.5	1037	2	US-09-057-121-1	Sequence 1, Appli
	678	15	0.5	1037	4	US-09-358-734-1	Sequence 1, Appli
	679	15	0.5	1051	4	US-09-245-041-10	Sequence 10, Appl
	680	15	0.5	1056	4	US-09-524-162-1	Sequence 1, Appli
	681	15	0.5	1058	4	US-09-452-239-11	Sequence 11, Appl
c	682	15	0.5	1068	4	US-09-403-768-5	Sequence 5, Appli
c	683	15	0.5	1070	4	US-09-470-443-7	Sequence 7, Appli
	684	15	0.5	1083	2	US-08-319-866-1	Sequence 1, Appli

685	15	0.5	1104	4	US-09-009-816-1	Sequence 1, Appli
c 686	15	0.5	1104	4	US-09-009-816-1	Sequence 1, Appli
687	15	0.5	1112	4	US-09-173-300-48	Sequence 48, Appl
688	15	0.5	1116	4	US-08-776-971-139	Sequence 139, App
689	15	0.5	1127	4	US-09-342-084-1	Sequence 1, Appli
690	15	0.5	1130	4	US-09-221-017B-289	Sequence 289, App
c 691	15	0.5	1137	1	US-08-706-214-2	Sequence 2, Appli
692	15	0.5	1144	1	US-08-014-943A-1	Sequence 1, Appli
693	15	0.5	1144	1	US-08-486-421-2	Sequence 2, Appli
694	15	0.5	1144	1	US-08-470-911-2	Sequence 2, Appli
695	15	0.5	1144	2	US-08-486-809-2	Sequence 2, Appli
696	15	0.5	1147	1	US-08-417-103-15	Sequence 15, Appl
697	15	0.5	1151	4	US-09-430-503-1	Sequence 1, Appli
698	15	0.5	1151	4	US-09-430-503-3	Sequence 3, Appli
699	15	0.5	1151	4	US-09-430-503-5	Sequence 5, Appli
700	15	0.5	1151	4	US-09-430-503-7	Sequence 7, Appli
701	15	0.5	1154	3	US-08-651-136C-7	Sequence 7, Appli
702	15	0.5	1154	4	US-09-229-911A-7	Sequence 7, Appli
703	15	0.5	1167	2	US-08-492-027A-5	Sequence 5, Appli
c 704	15	0.5	1167	2	US-08-492-027A-5	Sequence 5, Appli
705	15	0.5	1173	3	US-08-993-380-3	Sequence 3, Appli
c 706	15	0.5	1176	4	US-09-372-422A-25	Sequence 25, Appl
707	15	0.5	1177	1	US-08-249-671A-8	Sequence 8, Appli
708	15	0.5	1185	2	US-08-628-039-9	Sequence 9, Appli
709	15	0.5	1185	3	US-08-912-205-9	Sequence 9, Appli
710	15	0.5	1185	4	US-09-440-400-9	Sequence 9, Appli
c 711	15	0.5	1191	4	US-09-134-001C-1709	Sequence 1709, Ap
712	15	0.5	1200	4	US-09-222-938A-47	Sequence 47, Appl
713	15	0.5	1202	4	US-09-293-322C-9	Sequence 9, Appli
714	15	0.5	1203	1	US-08-362-670B-29	Sequence 29, Appl
715	15	0.5	1203	3	US-08-333-576C-29	Sequence 29, Appl
716	15	0.5	1203	4	US-09-086-010-1	Sequence 1, Appli
717	15	0.5	1203	4	US-08-808-324-29	Sequence 29, Appl
718	15	0.5	1203	5	PCT-US94-14030A-29	Sequence 29, Appl
c 719	15	0.5	1206	1	US-08-318-947A-1	Sequence 1, Appli
c 720	15	0.5	1206	2	US-08-795-303-1	Sequence 1, Appli
c 721	15	0.5	1213	1	US-08-276-151-6	Sequence 6, Appli
c 722	15	0.5	1223	4	US-08-957-351-4	Sequence 4, Appli
c 723	15	0.5	1223	4	US-08-957-351-29	Sequence 29, Appl
724	15	0.5	1233	4	US-08-765-907A-12	Sequence 12, Appl
c 725	15	0.5	1240	4	US-08-957-351-8	Sequence 8, Appli
c 726	15	0.5	1242	2	US-08-960-756-1	Sequence 1, Appli
727	15	0.5	1244	1	US-07-816-283-7	Sequence 7, Appli
728	15	0.5	1244	1	US-08-417-103-7	Sequence 7, Appli
c 729	15	0.5	1261	4	US-09-268-140-10	Sequence 10, Appl
c 730	15	0.5	1280	4	US-08-483-533-38	Sequence 38, Appl
c 731	15	0.5	1280	4	US-09-283-471A-38	Sequence 38, Appl
732	15	0.5	1284	3	US-08-985-526-24	Sequence 24, Appl
733	15	0.5	1287	1	US-08-785-050-1	Sequence 1, Appli
734	15	0.5	1287	2	US-08-996-798-1	Sequence 1, Appli
c 735	15	0.5	1287	4	US-09-134-001C-48	Sequence 48, Appl
c 736	15	0.5	1300	4	US-08-483-533-39	Sequence 39, Appl
c 737	15	0.5	1300	4	US-09-283-471A-39	Sequence 39, Appl
c 738	15	0.5	1302	4	US-09-372-422A-27	Sequence 27, Appl
739	15	0.5	1312	1	US-08-205-506A-1	Sequence 1, Appli
c 740	15	0.5	1312	1	US-08-205-506A-1	Sequence 1, Appli
741	15	0.5	1312	5	PCT-US94-02389-1	Sequence 1, Appli
c 742	15	0.5	1312	5	PCT-US94-02389-1	Sequence 1, Appli
743	15	0.5	1315	2	US-08-381-881-1	Sequence 1, Appli
744	15	0.5	1315	4	US-09-281-221-1	Sequence 1, Appli
745	15	0.5	1318	3	US-08-491-954-1	Sequence 1, Appli
c 746	15	0.5	1332	2	US-09-057-762-1	Sequence 1, Appli
c 747	15	0.5	1332	3	US-08-326-119A-1	Sequence 1, Appli
748	15	0.5	1335	2	US-08-985-090-3	Sequence 3, Appli
749	15	0.5	1335	3	US-09-165-543-3	Sequence 3, Appli
750	15	0.5	1349	4	US-09-191-608-15	Sequence 15, Appl
751	15	0.5	1351	1	US-07-816-283-5	Sequence 5, Appli
752	15	0.5	1351	1	US-08-417-103-5	Sequence 5, Appli
753	15	0.5	1352	1	US-07-714-819-1	Sequence 1, Appli
754	15	0.5	1352	1	US-08-246-978A-1	Sequence 1, Appli
755	15	0.5	1352	1	US-08-440-814A-1	Sequence 1, Appli
756	15	0.5	1368	4	US-09-035-382-1	Sequence 1, Appli
c 757	15	0.5	1380	2	US-08-748-947A-1	Sequence 1, Appli

758	15	0.5	1388	4	US-09-008-271A-22	Sequence 22, Appl
c 759	15	0.5	1389	4	US-09-134-001C-1289	Sequence 1289, Ap
760	15	0.5	1404	4	US-09-257-179-34	Sequence 34, Appl
c 761	15	0.5	1405	1	US-08-035-558-3	Sequence 3, Appli
c 762	15	0.5	1405	2	US-08-682-847-5	Sequence 5, Appli
c 763	15	0.5	1405	4	US-09-063-676-1	Sequence 1, Appli
c 764	15	0.5	1410	1	US-08-353-550-5	Sequence 5, Appli
c 765	15	0.5	1410	2	US-08-551-687-5	Sequence 5, Appli
766	15	0.5	1419	4	US-09-516-914-4	Sequence 4, Appli
c 767	15	0.5	1420	4	US-09-482-180A-1	Sequence 1, Appli
768	15	0.5	1421	4	US-09-191-608-14	Sequence 14, Appl
c 769	15	0.5	1433	2	US-08-666-392A-1	Sequence 1, Appli
c 770	15	0.5	1433	4	US-09-199-926-1	Sequence 1, Appli
c 771	15	0.5	1435	3	US-09-143-212-1	Sequence 1, Appli
772	15	0.5	1436	4	US-09-191-608-13	Sequence 13, Appl
c 773	15	0.5	1441	1	US-08-414-625-1	Sequence 1, Appli
774	15	0.5	1443	4	US-09-537-168-1	Sequence 1, Appli
c 775	15	0.5	1459	4	US-09-180-109A-13	Sequence 13, Appl
776	15	0.5	1467	1	US-08-176-620A-3	Sequence 3, Appli
777	15	0.5	1467	1	US-08-463-862-3	Sequence 3, Appli
778	15	0.5	1467	2	US-08-461-985-3	Sequence 3, Appli
779	15	0.5	1467	2	US-08-458-887-3	Sequence 3, Appli
780	15	0.5	1467	4	US-08-932-787B-3	Sequence 3, Appli
781	15	0.5	1467	4	US-08-932-012C-3	Sequence 3, Appli
782	15	0.5	1467	4	US-08-888-818C-3	Sequence 3, Appli
c 783	15	0.5	1479	6	5486473-5	Patent No. 5486473
784	15	0.5	1499	4	US-09-191-608-16	Sequence 16, Appl
785	15	0.5	1500	1	US-08-443-568B-15	Sequence 15, Appl
786	15	0.5	1500	5	PCT-US94-06997-15	Sequence 15, Appl
787	15	0.5	1515	4	US-09-513-729B-3	Sequence 3, Appli
c 788	15	0.5	1515	4	US-09-513-729B-3	Sequence 3, Appli
c 789	15	0.5	1523	1	US-07-795-859B-25	Sequence 25, Appl
c 790	15	0.5	1523	1	US-08-457-616-25	Sequence 25, Appl
791	15	0.5	1528	4	US-09-362-473-13	Sequence 13, Appl
792	15	0.5	1530	4	US-09-444-336-3	Sequence 3, Appli
c 793	15	0.5	1536	4	US-09-134-001C-2579	Sequence 2579, Ap
794	15	0.5	1537	4	US-09-449-285A-10	Sequence 10, Appl
795	15	0.5	1537	4	US-09-149-476-311	Sequence 311, App
796	15	0.5	1555	2	US-08-696-376-1	Sequence 1, Appli
c 797	15	0.5	1556	4	US-09-043-937A-3	Sequence 3, Appli
798	15	0.5	1561	4	US-08-765-907A-11	Sequence 11, Appl
c 799	15	0.5	1567	1	US-08-722-001-24	Sequence 24, Appl
c 800	15	0.5	1570	3	US-08-617-860B-20	Sequence 20, Appl
c 801	15	0.5	1591	1	US-07-865-878A-3	Sequence 3, Appli
c 802	15	0.5	1591	1	US-07-676-647-1	Sequence 1, Appli
c 803	15	0.5	1591	1	US-08-449-329-1	Sequence 1, Appli
c 804	15	0.5	1591	2	US-08-445-073-1	Sequence 1, Appli
c 805	15	0.5	1591	2	US-08-585-258-1	Sequence 1, Appli
c 806	15	0.5	1591	2	US-08-603-010-3	Sequence 3, Appli
c 807	15	0.5	1591	4	US-09-211-590-1	Sequence 1, Appli
c 808	15	0.5	1591	5	PCT-US91-03896-1	Sequence 1, Appli
809	15	0.5	1605	4	US-09-124-541-3	Sequence 3, Appli
810	15	0.5	1605	4	US-09-149-476-187	Sequence 187, App
811	15	0.5	1611	3	US-08-909-742-2	Sequence 2, Appli
812	15	0.5	1611	4	US-09-412-289-2	Sequence 2, Appli
813	15	0.5	1617	4	US-09-444-336-4	Sequence 4, Appli
814	15	0.5	1621	4	US-08-679-645-1	Sequence 1, Appli
815	15	0.5	1631	4	US-09-149-476-97	Sequence 97, Appl
816	15	0.5	1632	1	US-08-324-243-34	Sequence 34, Appl
817	15	0.5	1632	1	US-08-532-390-34	Sequence 34, Appl
818	15	0.5	1632	3	US-08-717-294-34	Sequence 34, Appl
819	15	0.5	1632	5	PCT-US95-11511-34	Sequence 34, Appl
820	15	0.5	1634	4	US-09-149-476-291	Sequence 291, App
821	15	0.5	1635	4	US-09-178-252-10	Sequence 10, Appl
c 822	15	0.5	1638	4	US-09-351-224E-2	Sequence 2, Appli
823	15	0.5	1659	3	US-09-083-351-3	Sequence 3, Appli
824	15	0.5	1659	4	US-09-083-352-3	Sequence 3, Appli
c 825	15	0.5	1662	1	US-08-336-408B-5	Sequence 5, Appli
c 826	15	0.5	1662	5	PCT-US91-00399-5	Sequence 5, Appli
c 827	15	0.5	1680	1	US-08-234-783-3	Sequence 3, Appli
c 828	15	0.5	1680	1	US-08-456-907-3	Sequence 3, Appli
c 829	15	0.5	1680	5	PCT-US95-05523-3	Sequence 3, Appli
c 830	15	0.5	1691	4	US-09-351-224E-1	Sequence 1, Appli

c	831	15	0.5	1700	5	PCT-US95-04801-1	Sequence 1, Appli
c	832	15	0.5	1701	4	US-09-232-468A-21	Sequence 21, Appl
	833	15	0.5	1727	3	US-08-999-733-2	Sequence 2, Appli
	834	15	0.5	1730	4	US-09-575-574-5	Sequence 5, Appli
c	835	15	0.5	1739	4	US-09-594-193-1	Sequence 1, Appli
c	836	15	0.5	1741	1	US-08-565-655-5	Sequence 5, Appli
	837	15	0.5	1752	4	US-09-152-060-38	Sequence 38, Appl
	838	15	0.5	1762	2	US-08-742-621-2	Sequence 2, Appli
	839	15	0.5	1769	2	US-08-765-268A-1	Sequence 1, Appli
c	840	15	0.5	1770	3	US-09-006-636-1	Sequence 1, Appli
c	841	15	0.5	1770	4	US-09-006-632-1	Sequence 1, Appli
c	842	15	0.5	1770	4	US-09-325-274-1	Sequence 1, Appli
	843	15	0.5	1776	1	US-08-722-001-29	Sequence 29, Appl
c	844	15	0.5	1778	4	US-08-934-386-4	Sequence 4, Appli
	845	15	0.5	1791	4	US-09-134-001C-526	Sequence 526, App
	846	15	0.5	1797	4	US-09-364-230-19	Sequence 19, Appl
	847	15	0.5	1821	4	US-09-486-382B-10	Sequence 10, Appl
c	848	15	0.5	1834	4	US-09-105-390-4	Sequence 4, Appli
c	849	15	0.5	1835	1	US-08-361-611-3	Sequence 3, Appli
c	850	15	0.5	1835	1	US-08-565-655-3	Sequence 3, Appli
c	851	15	0.5	1835	2	US-08-946-967-3	Sequence 3, Appli
	852	15	0.5	1841	4	US-09-053-374A-1	Sequence 1, Appli
c	853	15	0.5	1851	4	US-09-183-959-9	Sequence 9, Appli
	854	15	0.5	1857	4	US-09-377-557-9	Sequence 9, Appli
	855	15	0.5	1866	3	US-08-909-742-1	Sequence 1, Appli
	856	15	0.5	1866	4	US-09-412-289-1	Sequence 1, Appli
c	857	15	0.5	1896	4	US-09-343-011B-4	Sequence 4, Appli
	858	15	0.5	1897	1	US-08-453-477-1	Sequence 1, Appli
	859	15	0.5	1897	1	US-08-453-461-1	Sequence 1, Appli
c	860	15	0.5	1920	4	US-09-160-496-4	Sequence 4, Appli
	861	15	0.5	1946	2	US-08-755-584-1	Sequence 1, Appli
	862	15	0.5	1946	3	US-09-192-611-1	Sequence 1, Appli
c	863	15	0.5	1962	1	US-08-399-696-1	Sequence 1, Appli
	864	15	0.5	1984	4	US-08-937-067-5	Sequence 5, Appli
	865	15	0.5	1995	3	US-08-904-452-1	Sequence 1, Appli
	866	15	0.5	1995	4	US-09-517-639-1	Sequence 1, Appli
c	867	15	0.5	2018	1	US-08-188-582-6	Sequence 6, Appli
c	868	15	0.5	2018	1	US-08-646-715-6	Sequence 6, Appli
	869	15	0.5	2025	3	US-09-160-483-2	Sequence 2, Appli
	870	15	0.5	2031	4	US-09-149-476-155	Sequence 155, App
	871	15	0.5	2033	1	US-08-148-910-14	Sequence 14, Appl
	872	15	0.5	2033	1	US-08-448-937A-14	Sequence 14, Appl
	873	15	0.5	2084	1	US-08-720-899-3	Sequence 3, Appli
	874	15	0.5	2084	1	US-08-459-610-3	Sequence 3, Appli
	875	15	0.5	2084	2	US-08-343-804-3	Sequence 3, Appli
	876	15	0.5	2084	2	US-08-687-399-3	Sequence 3, Appli
	877	15	0.5	2084	2	US-08-600-908A-3	Sequence 3, Appli
	878	15	0.5	2084	3	US-08-683-838A-3	Sequence 3, Appli
	879	15	0.5	2084	3	US-09-182-859-3	Sequence 3, Appli
	880	15	0.5	2084	4	US-09-170-670-13	Sequence 13, Appl
	881	15	0.5	2084	4	US-09-193-068-31	Sequence 31, Appl
	882	15	0.5	2084	4	US-09-183-412-58	Sequence 58, Appl
	883	15	0.5	2084	4	US-09-290-734-13	Sequence 13, Appl
	884	15	0.5	2084	4	US-09-672-459-3	Sequence 3, Appli
	885	15	0.5	2084	4	US-09-636-252A-3	Sequence 3, Appli
c	886	15	0.5	2085	4	US-09-352-990-9	Sequence 9, Appli
	887	15	0.5	2087	4	US-09-097-199-83	Sequence 83, Appl
	888	15	0.5	2094	4	US-09-106-194-1	Sequence 1, Appli
c	889	15	0.5	2094	4	US-09-661-753-1	Sequence 1, Appli
c	890	15	0.5	2108	4	US-09-032-742-6	Sequence 6, Appli
	891	15	0.5	2114	4	US-09-568-816A-1	Sequence 1, Appli
	892	15	0.5	2115	2	US-08-474-379C-60	Sequence 60, Appl
c	893	15	0.5	2115	2	US-08-474-379C-60	Sequence 60, Appl
	894	15	0.5	2115	3	US-09-146-249A-60	Sequence 60, Appl
c	895	15	0.5	2115	3	US-09-146-249A-60	Sequence 60, Appl
	896	15	0.5	2115	3	US-08-206-188B-60	Sequence 60, Appl
c	897	15	0.5	2115	3	US-08-206-188B-60	Sequence 60, Appl
	898	15	0.5	2130	1	US-07-952-800-1	Sequence 1, Appli
c	899	15	0.5	2135	4	US-09-173-300-1	Sequence 1, Appli
	900	15	0.5	2140	1	US-08-334-698-1	Sequence 1, Appli
	901	15	0.5	2140	1	US-08-228-932-1	Sequence 1, Appli
	902	15	0.5	2140	1	US-08-468-939-1	Sequence 1, Appli
	903	15	0.5	2140	2	US-08-406-855A-1	Sequence 1, Appli
	904	15	0.5	2140	2	US-08-722-190-1	Sequence 1, Appli
	905	15	0.5	2140	3	US-08-244-354-1	Sequence 1, Appli
	906	15	0.5	2140	3	US-09-206-899-1	Sequence 1, Appli
	907	15	0.5	2140	4	US-09-444-783-1	Sequence 1, Appli
	908	15	0.5	2140	4	US-09-688-415-1	Sequence 1, Appli
	909	15	0.5	2140	5	PCT-US95-04203-1	Sequence 1, Appli
	910	15	0.5	2143	3	US-09-097-309-5	Sequence 5, Appli
	911	15	0.5	2143	4	US-09-097-171A-9	Sequence 9, Appli
	912	15	0.5	2143	4	US-09-460-587-5	Sequence 5, Appli
	913	15	0.5	2156	4	US-08-965-762-16	Sequence 16, Appl
	914	15	0.5	2156	4	US-09-911-927-16	Sequence 16, Appl
c	915	15	0.5	2156	4	US-09-911-927-18	Sequence 18, Appl
	916	15	0.5	2156	4	US-09-911-882-16	Sequence 16, Appl
c	917	15	0.5	2156	4	US-09-911-882-18	Sequence 18, Appl
	918	15	0.5	2178	1	US-08-463-587A-24	Sequence 24, Appl
	919	15	0.5	2178	2	US-08-463-667A-2	Sequence 2, Appli
	920	15	0.5	2178	3	US-08-923-854-24	Sequence 24, Appl
	921	15	0.5	2178	5	PCT-US91-09133-25	Sequence 25, Appl
c	922	15	0.5	2262	1	US-07-689-008-3	Sequence 3, Appli
	923	15	0.5	2308	4	US-09-382-256-9	Sequence 9, Appli
	924	15	0.5	2308	4	US-09-395-115-9	Sequence 9, Appli
	925	15	0.5	2308	4	US-08-436-265-9	Sequence 9, Appli
	926	15	0.5	2308	4	US-09-679-187-9	Sequence 9, Appli
c	927	15	0.5	2369	4	US-09-302-769-20	Sequence 20, Appl
	928	15	0.5	2398	1	US-07-971-558-2	Sequence 2, Appli
	929	15	0.5	2398	2	US-08-515-251A-3	Sequence 3, Appli
c	930	15	0.5	2412	1	US-08-158-232-9	Sequence 9, Appli
c	931	15	0.5	2412	1	US-08-304-626-9	Sequence 9, Appli
c	932	15	0.5	2412	1	US-08-316-301A-11	Sequence 11, Appl
c	933	15	0.5	2412	1	US-08-611-928-9	Sequence 9, Appli
c	934	15	0.5	2412	3	US-09-173-891-9	Sequence 9, Appli
c	935	15	0.5	2412	4	US-09-076-137-11	Sequence 11, Appl
c	936	15	0.5	2412	5	PCT-US92-03624-11	Sequence 11, Appl
	937	15	0.5	2419	4	US-09-245-041-8	Sequence 8, Appli
c	938	15	0.5	2461	1	US-08-282-141-1	Sequence 1, Appli
	939	15	0.5	2481	1	US-08-324-243-35	Sequence 35, Appl
	940	15	0.5	2481	1	US-08-532-390-35	Sequence 35, Appl
	941	15	0.5	2481	3	US-08-717-294-35	Sequence 35, Appl
	942	15	0.5	2481	5	PCT-US95-11511-35	Sequence 35, Appl
	943	15	0.5	2505	4	US-09-097-199-85	Sequence 85, Appl
c	944	15	0.5	2505	4	US-09-268-140-1	Sequence 1, Appli
c	945	15	0.5	2510	4	US-08-894-324A-2	Sequence 2, Appli
	946	15	0.5	2511	1	US-08-363-560-1	Sequence 1, Appli
c	947	15	0.5	2517	4	US-09-268-140-7	Sequence 7, Appli
	948	15	0.5	2538	3	US-08-899-437-1	Sequence 1, Appli
	949	15	0.5	2538	4	US-09-126-121-1	Sequence 1, Appli
c	950	15	0.5	2544	4	US-09-483-371-1	Sequence 1, Appli
c	951	15	0.5	2565	4	US-09-105-058C-26	Sequence 26, Appl
	952	15	0.5	2604	4	US-09-264-097-3	Sequence 3, Appli
	953	15	0.5	2604	4	US-09-537-168-5	Sequence 5, Appli
	954	15	0.5	2609	4	US-09-221-017B-747	Sequence 747, App
c	955	15	0.5	2629	2	US-09-012-030-1	Sequence 1, Appli
c	956	15	0.5	2629	2	US-08-590-454-1	Sequence 1, Appli
c	957	15	0.5	2639	4	US-09-374-135-1	Sequence 1, Appli
c	958	15	0.5	2653	1	US-08-295-670-5	Sequence 5, Appli
c	959	15	0.5	2653	1	US-08-633-485-5	Sequence 5, Appli
	960	15	0.5	2677	4	US-08-882-164D-36	Sequence 36, Appl
	961	15	0.5	2689	2	US-08-985-090-1	Sequence 1, Appli
	962	15	0.5	2689	3	US-09-165-543-1	Sequence 1, Appli
	963	15	0.5	2693	1	US-07-872-644-50	Sequence 50, Appl
c	964	15	0.5	2693	1	US-08-295-670-7	Sequence 7, Appli
	965	15	0.5	2693	1	US-08-297-494-50	Sequence 50, Appl
	966	15	0.5	2693	1	US-08-297-510-50	Sequence 50, Appl
c	967	15	0.5	2693	1	US-08-633-485-7	Sequence 7, Appli
	968	15	0.5	2693	1	US-08-479-532-50	Sequence 50, Appl
	969	15	0.5	2693	1	US-08-455-526-50	Sequence 50, Appl
	970	15	0.5	2693	1	US-08-455-525-50	Sequence 50, Appl
	971	15	0.5	2693	3	US-09-139-491-50	Sequence 50, Appl
	972	15	0.5	2693	5	PCT-US92-03222-50	Sequence 50, Appl
c	973	15	0.5	2702	3	US-08-508-761B-3	Sequence 3, Appli
c	974	15	0.5	2742	4	US-09-232-468A-1	Sequence 1, Appli
	975	15	0.5	2757	2	US-08-627-254C-26	Sequence 26, Appl
c	976	15	0.5	2764	2	US-08-465-971B-1	Sequence 1, Appli

```
c 977     15    0.5    2787  4  US-09-624-693A-16        Sequence 16, Appl
  978     15    0.5    2817  5  PCT-US93-05944-1         Sequence 1, Appli
  979     15    0.5    2838  4  US-08-246-489-1          Sequence 1, Appli
  980     15    0.5    2853  4  US-09-221-017B-843       Sequence 843, App
  981     15    0.5    2870  1  US-08-468-036-28         Sequence 28, Appl
  982     15    0.5    2870  2  US-08-376-843-28         Sequence 28, Appl
  983     15    0.5    2892  1  US-08-264-534-5          Sequence 5, Appli
  984     15    0.5    2892  1  US-08-083-590A-1         Sequence 1, Appli
  985     15    0.5    2892  1  US-08-465-500-5          Sequence 5, Appli
  986     15    0.5    2892  2  US-08-346-126-5          Sequence 5, Appli
  987     15    0.5    2892  2  US-08-346-128-5          Sequence 5, Appli
  988     15    0.5    2892  3  US-08-532-384-1          Sequence 1, Appli
  989     15    0.5    2892  3  US-08-893-828-5          Sequence 5, Appli
  990     15    0.5    2897  2  US-08-927-394-1          Sequence 1, Appli
c 991     15    0.5    2908  4  US-09-930-181-1          Sequence 1, Appli
c 992     15    0.5    2914  4  US-09-177-650-6          Sequence 6, Appli
c 993     15    0.5    2991  3  US-08-795-430-48         Sequence 48, Appl
c 994     15    0.5    2991  4  US-09-355-700-48         Sequence 48, Appl
  995     15    0.5    3000  4  US-08-460-269C-3         Sequence 3, Appli
  996     15    0.5    3000  4  US-08-460-269C-5         Sequence 5, Appli
c 997     15    0.5    3044  1  US-07-872-644-44         Sequence 44, Appl
c 998     15    0.5    3044  1  US-08-297-494-44         Sequence 44, Appl
c 999     15    0.5    3044  1  US-08-297-510-44         Sequence 44, Appl
c1000     15    0.5    3044  1  US-08-479-532-44         Sequence 44, Appl
```

ALIGNMENTS

```
RESULT 1
US-08-750-145A-23/c
; Sequence 23, Application US/08750145A
; Patent No. 6010851
;  GENERAL INFORMATION:
;    APPLICANT:  MIHARA, Yasuhiro
;    APPLICANT:  UTAGAWA, Takashi
;    APPLICANT:  YAMADA, Hideaki
;    APPLICANT:  ASANO, Yasuhisa
;    TITLE OF INVENTION:  Method for Producing Nucleoside-5'-
;    TITLE OF INVENTION:  Phosphate Ester
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
;      STREET:  1755 S. JEFF. DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/750,145A
;      FILING DATE:  01-JAN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-149781
;      FILING DATE:  05-May-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-094680
;      FILING DATE:  26-Mar-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORMAN F. OBLON
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  0010-0830-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Serratia ficaria
;      STRAIN:  IAM 13540
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..732
US-08-750-145A-23

  Query Match             0.6%;  Score 20;  DB 3;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1751 TGCAGCTGGGCGGCGAAGGT 1770
          ||||||||||||||||||||
Db    698 TGCAGCTGGGCGGCGAAGGT 679

RESULT 2
US-08-975-698A-27/c
; Sequence 27, Application US/08975698A
; Patent No. 6015697
;  GENERAL INFORMATION:
;    APPLICANT:  MIHARA, YASUHIRO
;    APPLICANT:  UTAGAWA, TAKASHI
;    APPLICANT:  YAMADA, HIDEAKI
;    APPLICANT:  ASANO, YASUHISA
;    TITLE OF INVENTION:  METHOD FOR PRODUCING NUCLEOSIDE-5'-PHOSPHATE
;    TITLE OF INVENTION:  ESTER
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,698A
;      FILING DATE:  21-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  0010-0885-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Serratia ficaria
;      STRAIN:  IAM 13540
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..732
```

```
US-08-975-698A-27

  Query Match             0.6%;  Score 20;  DB 3;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1751 TGCAGCTGGGCGGCGAAGGT 1770
          ||||||||||||||||||||
Db    698 TGCAGCTGGGCGGCGAAGGT 679


RESULT 3
US-09-417-090-27/c
; Sequence 27, Application US/09417090
; Patent No. 6207435
;  GENERAL INFORMATION:
;       APPLICANT: MIHARA, YASUHIRO
;                  UTAGAWA, TAKASHI
;                  YAMADA, HIDEAKI
;                  ASANO, YASUHISA
;       TITLE OF INVENTION: METHOD FOR PRODUCING NUCLEOSIDE-5'-PHOSPHATE
;                           ESTER
;       NUMBER OF SEQUENCES: 30
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
;            STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
;            CITY: ARLINGTON
;            STATE: VIRGINIA
;            COUNTRY: USA
;            ZIP: 22202
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/417,090
;            FILING DATE: 13-Oct-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/975,698
;            FILING DATE: 21-NOV-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: OBLON, NORMAN F
;            REGISTRATION NUMBER: 24,618
;            REFERENCE/DOCKET NUMBER: 0010-0885-0
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (703) 413-3000
;            TELEFAX: (703) 413-2220
;  INFORMATION FOR SEQ ID NO: 27:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 735 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: genomic DNA
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;            ORGANISM: Serratia ficaria
;            STRAIN: IAM 13540
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..732
;       SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-417-090-27

  Query Match             0.6%;  Score 20;  DB 4;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1751 TGCAGCTGGGCGGCGAAGGT 1770
          ||||||||||||||||||||
Db    698 TGCAGCTGGGCGGCGAAGGT 679


RESULT 4
US-09-727-578-27/c
; Sequence 27, Application US/09727578
; Patent No. 6355472
;  GENERAL INFORMATION:
;    APPLICANT:  MIHARA, YASUHIRO
;    APPLICANT:  UTAGAWA, TAKASHI
;    APPLICANT:  YAMADA, HIDEAKI
;    APPLICANT:  ASANO, YASUHISA
;    TITLE OF INVENTION:  METHOD FOR PRODUCING NUCLEOSIDE-5'-PHOSPHATE
;    TITLE OF INVENTION:  ESTER
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/727,578
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/975,698
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  0010-0885-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Serratia ficaria
;      STRAIN:  IAM 13540
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..732
US-09-727-578-27

  Query Match             0.6%;  Score 20;  DB 4;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1751 TGCAGCTGGGCGGCGAAGGT 1770
          ||||||||||||||||||||
Db    698 TGCAGCTGGGCGGCGAAGGT 679


RESULT 5
US-08-219-012-74/c
; Sequence 74, Application US/08219012
; Patent No. 5543293
;  GENERAL INFORMATION:
;    APPLICANT:  Larry Gold
```

```
;    APPLICANT:  Diane Tasset
;    TITLE OF INVENTION:  Ligands of Thrombin
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Beaton & Swanson, P.C.
;      STREET:  4582 South Ulster Street Parkway, Suite #
;      STREET:  403
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80237
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 Kb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/219,012
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:  none
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 850-9900
;      TELEFAX:  (303) 850-9401
;  INFORMATION FOR SEQ ID NO:  74:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-219-012-74

  Query Match             0.6%;  Score 18;  DB 1;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        9 GACCCCAACCTTCCCTAC 26
            ||||||||||||||||||
Db       41 GACCCCAACCTTCCCTAC 24


RESULT 6
US-08-687-421-262/c
; Sequence 262, Application US/08687421
; Patent No. 6177557
;  GENERAL INFORMATION:
;    APPLICANT:  Gold, Larry
;    APPLICANT:  Janjic, Nebojsa
;    APPLICANT:  Tasset, Diane
;    TITLE OF INVENTION:  HIGH-AFFINITY LIGANDS OF BASIC
;    TITLE OF INVENTION:  FIBROBLAST GROWTH FACTOR AND
;    TITLE OF INVENTION:  THROMBIN
;    NUMBER OF SEQUENCES:  445
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MB storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/687,421
;      FILING DATE:  08-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/195,005
;      FILING DATE:  10-FEBRUARY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:  22-APRIL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/219,012
;      FILING DATE:  28-MARCH-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/973,333
;      FILING DATE:  11-NOVEMBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX07/PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  262:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-687-421-262

  Query Match             0.6%;  Score 18;  DB 4;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        9 GACCCCAACCTTCCCTAC 26
            ||||||||||||||||||
Db       41 GACCCCAACCTTCCCTAC 24


RESULT 7
US-09-221-017B-718/c
; Sequence 718, Application US/09221017B
; Patent No. 6444799
;  GENERAL INFORMATION:
;    APPLICANT:  Ross, Bruce C.
;    TITLE OF INVENTION:  P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
;    NUMBER OF SEQUENCES:  1120
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/221,017B
;      FILING DATE:  23-DEC-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP1182
;      FILING DATE:  31-DEC-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP1546
;      FILING DATE:  30-JAN-1998
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  PP2911
;      FILING DATE:  09-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU98/01023
;      FILING DATE:  10-DEC-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monroy, Gladys H
;      REGISTRATION NUMBER:  32,430
;      REFERENCE/DOCKET NUMBER:  27340-20021.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  718:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  692 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  UNKNOWN
;    ORIGINAL SOURCE:
;      ORGANISM:  PORYPHYROMONAS GINGIVALIS
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:   1...692
US-09-221-017B-718

  Query Match              0.6%;  Score 18;  DB 4;  Length 692;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2118 CAATCTGGAAAACCTGAT 2135
          ||||||||||||||||||
Db    561 CAATCTGGAAAACCTGAT 544


RESULT 8
US-09-276-531-101/c
; Sequence 101, Application US/09276531
; Patent No. 6183968
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Reddy, Roopa
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF GENES ENCODING
;    TITLE OF INVENTION:  RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
;    NUMBER OF SEQUENCES:  134
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/276,531
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/079,677
;      FILING DATE:  March 27, 1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lynn E. Murry, Ph.D.
;      REGISTRATION NUMBER:  42,918
;      REFERENCE/DOCKET NUMBER:  PA-0008 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  101:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1689 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BLADTUT07
;      CLONE:  1887573
US-09-276-531-101

  Query Match              0.6%;  Score 18;  DB 4;  Length 1689;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2541 GGGCATGGGACACAGCAC 2558
          ||||||||||||||||||
Db   1275 GGGCATGGGACACAGCAC 1258


RESULT 9
US-08-248-474-87
; Sequence 87, Application US/08248474
; Patent No. 5612471
;  GENERAL INFORMATION:
;    APPLICANT:  McK. BIRD, David
;    APPLICANT:  WILSON, Mark A.
;    TITLE OF INVENTION:  NEMATODE-INDUCE GENES IN TOMATO
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/248,474
;      FILING DATE:  25-MAY-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  2307E-535
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  87:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Lycopersicon esculentum cv 'Rutgers Large
;      ORGANISM:  Red'
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..77
;      OTHER INFORMATION:  /standard_name= "DB# 239"
```

```
US-08-248-474-87

  Query Match             0.5%;  Score 17;  DB 1;  Length 77;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db       6 AGGCGGCGGCGGCGGCA 22


RESULT 10
US-08-756-849-87
; Sequence 87, Application US/08756849
; Patent No. 6093810
;  GENERAL INFORMATION:
;    APPLICANT:  Bird, David McK.
;    APPLICANT:  Wilson, Mark A.
;    TITLE OF INVENTION:  Nematode-Induced Genes in Tomato
;    NUMBER OF SEQUENCES:  129
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/756,849
;      FILING DATE:  26-NOV-1996
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/248,474
;      FILING DATE:  25-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  023070-053510US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  87:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Lycopersicon esculentum cv 'Rutgers Large Red'
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..77
;      OTHER INFORMATION:  /standard_name= "DB# 239"
US-08-756-849-87

  Query Match             0.5%;  Score 17;  DB 3;  Length 77;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db       6 AGGCGGCGGCGGCGGCA 22


RESULT 11
US-08-906-156A-27
; Sequence 27, Application US/08906156A
; Patent No. 6287854
;  GENERAL INFORMATION:
;    APPLICANT:  SPURR, NIGEL K
;    APPLICANT:  GRAY, IAN C
;    APPLICANT:  STEWART, LORNA M
;    TITLE OF INVENTION:  DIAGNOSIS OF SUSCEPTIBILITY TO CANCER
;    TITLE OF INVENTION:  AND TREATMENT THEREOF
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/906,156A
;      FILING DATE:  05-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/042,655
;      FILING DATE:  02-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,147
;      FILING DATE:  13-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,840
;      FILING DATE:  23-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/96GB/02588
;      FILING DATE:  22-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SADOFF, B.J.
;      REGISTRATION NUMBER:  36,663
;      REFERENCE/DOCKET NUMBER:  1090-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  HUMAN
US-08-906-156A-27

  Query Match             0.5%;  Score 17;  DB 4;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      88 GGCGGCGGCGGCGGCAC 104
           |||||||||||||||||
Db      59 GGCGGCGGCGGCGGCAC 75


RESULT 12
US-08-906-156A-28
; Sequence 28, Application US/08906156A
; Patent No. 6287854
;  GENERAL INFORMATION:
;    APPLICANT:  SPURR, NIGEL K
;    APPLICANT:  GRAY, IAN C
;    APPLICANT:  STEWART, LORNA M
```

```
;    TITLE OF INVENTION:  DIAGNOSIS OF SUSCEPTIBILITY TO CANCER
;    TITLE OF INVENTION:  AND TREATMENT THEREOF
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/906,156A
;      FILING DATE:  05-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/042,655
;      FILING DATE:  02-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,147
;      FILING DATE:  13-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,840
;      FILING DATE:  23-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/96GB/02588
;      FILING DATE:  22-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SADOFF, B.J.
;      REGISTRATION NUMBER:  36,663
;      REFERENCE/DOCKET NUMBER:  1090-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  322 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  HUMAN
US-08-906-156A-28

  Query Match              0.5%;  Score 17;  DB 4;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      88 GGCGGCGGCGGCGGCAC 104
           |||||||||||||||||
Db      82 GGCGGCGGCGGCGGCAC 98

RESULT 13
US-09-641-638-222
; Sequence 222, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 222
;   LENGTH: 362
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 72
;   OTHER INFORMATION: 10-87-73 : polymorphic base C or T
;   NAME/KEY: misc_binding
;   LOCATION: 52..71
;   OTHER INFORMATION: 10-87-73.mis1, potential
;   NAME/KEY: misc_binding
;   LOCATION: 73..92
;   OTHER INFORMATION: 10-87-73.mis2, potential complement
;   NAME/KEY: primer_bind
;   LOCATION: 1..18
;   OTHER INFORMATION: upstream amplification primer
;   NAME/KEY: primer_bind
;   LOCATION: 345..362
;   OTHER INFORMATION: downstream amplification primer, complement
;   NAME/KEY: misc_binding
;   LOCATION: 60..84
;   OTHER INFORMATION: 10-87-73 potential probe
;   NAME/KEY: misc_feature
;   LOCATION: 172..174
;   OTHER INFORMATION: n=a, g, c or t
US-09-641-638-222

  Query Match              0.5%;  Score 17;  DB 4;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db      24 AGGCGGCGGCGGCGGCA 40

RESULT 14
US-09-641-638-223
; Sequence 223, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 223
```

```
;   LENGTH: 362
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 73
;   OTHER INFORMATION: 10-87-74 : polymorphic base A or T
;   NAME/KEY: misc_binding
;   LOCATION: 53..72
;   OTHER INFORMATION: 10-87-74.mis1, potential
;   NAME/KEY: misc_binding
;   LOCATION: 74..93
;   OTHER INFORMATION: 10-87-74.mis2, potential complement
;   NAME/KEY: primer_bind
;   LOCATION: 1..18
;   OTHER INFORMATION: upstream amplification primer
;   NAME/KEY: primer_bind
;   LOCATION: 345..362
;   OTHER INFORMATION: downstream amplification primer, complement
;   NAME/KEY: misc_binding
;   LOCATION: 61..85
;   OTHER INFORMATION: 10-87-74 potential probe
;   NAME/KEY: misc_feature
;   LOCATION: 172..174
;   OTHER INFORMATION: n=a, g, c or t
US-09-641-638-223

  Query Match             0.5%;  Score 17;  DB 4;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db      24 AGGCGGCGGCGGCGGCA 40


RESULT 15
US-09-641-638-224
; Sequence 224, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 224
;   LENGTH: 362
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 79
;   OTHER INFORMATION: 10-87-80 : polymorphic base A or G
;   NAME/KEY: misc_binding
;   LOCATION: 59..78
;   OTHER INFORMATION: 10-87-80.mis1, potential
;   NAME/KEY: misc_binding
;   LOCATION: 80..99
;   OTHER INFORMATION: 10-87-80.mis2, potential complement
;   NAME/KEY: primer_bind
;   LOCATION: 1..18
;   OTHER INFORMATION: upstream amplification primer
;   NAME/KEY: primer_bind
;   LOCATION: 345..362
;   OTHER INFORMATION: downstream amplification primer, complement
;   NAME/KEY: misc_binding
;   LOCATION: 67..91
;   OTHER INFORMATION: 10-87-80 potential probe
;   NAME/KEY: misc_feature
;   LOCATION: 172..174
;   OTHER INFORMATION: n=a, g, c or t
US-09-641-638-224

  Query Match             0.5%;  Score 17;  DB 4;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db      24 AGGCGGCGGCGGCGGCA 40


RESULT 16
US-09-641-638-225
; Sequence 225, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 225
;   LENGTH: 362
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 138
;   OTHER INFORMATION: 10-87-140 : polymorphic base C or T
;   NAME/KEY: misc_binding
;   LOCATION: 118..137
;   OTHER INFORMATION: 10-87-140.mis1, potential
;   NAME/KEY: misc_binding
;   LOCATION: 139..158
;   OTHER INFORMATION: 10-87-140.mis2, potential complement
;   NAME/KEY: primer_bind
;   LOCATION: 1..18
;   OTHER INFORMATION: upstream amplification primer
;   NAME/KEY: primer_bind
;   LOCATION: 345..362
;   OTHER INFORMATION: downstream amplification primer, complement
;   NAME/KEY: misc_binding
;   LOCATION: 126..150
;   OTHER INFORMATION: 10-87-140 potential probe
;   NAME/KEY: misc_feature
;   LOCATION: 172..174
```

```
;   OTHER INFORMATION: n=a, g, c or t
US-09-641-638-225

  Query Match              0.5%;  Score 17;  DB 4;  Length 362;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       87 AGGCGGCGGCGGCGGCA 103
            |||||||||||||||||
Db       24 AGGCGGCGGCGGCGGCA 40


RESULT 17
US-09-060-756-653
; Sequence 653, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
;  APPLICANT: Cole, Stewart
;  APPLICANT: Buchrieser-Brosch, Roland
;  APPLICANT: Gordon, Stephen
;  APPLICANT: Billault, Alain
;  TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
;  TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
;  TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
;  FILE REFERENCE: 3495-0169
;  CURRENT APPLICATION NUMBER: US/09/060,756
;  CURRENT FILING DATE: 1998-04-16
;  NUMBER OF SEQ ID NOS: 743
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 653
;   LENGTH: 429
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (various positions within the sequence)
;   OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-060-756-653

  Query Match              0.5%;  Score 17;  DB 4;  Length 429;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      792 TGGCGTGCGCATCGTCA 808
            |||||||||||||||||
Db      202 TGGCGTGCGCATCGTCA 218


RESULT 18
US-08-642-274D-18
; Sequence 18, Application US/08642274D
; Patent No. 6200749
; GENERAL INFORMATION:
;  APPLICANT:  Shiloh, Yosef
;  TITLE OF INVENTION:  MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO
;  TITLE OF INVENTION:  SCREEN FOR A PARTIAL A-T PHENOTYPE
;  FILE REFERENCE:  229000033
;  CURRENT APPLICATION NUMBER: US/08/642,274D
;  CURRENT FILING DATE:  1996-05-03
;  NUMBER OF SEQ ID NOS:  220
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 500
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: genomic
US-08-642-274D-18

  Query Match              0.5%;  Score 17;  DB 4;  Length 500;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      157 GCAACAACAGCGAAATC 173
            |||||||||||||||||
Db       43 GCAACAACAGCGAAATC 59


RESULT 19
US-08-952-014C-18
; Sequence 18, Application US/08952014C
; Patent No. 6265158
;  GENERAL INFORMATION:
;    APPLICANT:  Shiloh, Yosef
;    TITLE OF INVENTION:  ATAXIA-TELANGIECTASIA GENE AND ITS
;    TITLE OF INVENTION:  GENOMIC ORGANIZATION
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kohn & Associates
;      STREET:  30500 No. 6265158thwestern Hwy., Suite 410
;      CITY:  Farmington Hills
;      STATE:  Michigan
;      COUNTRY:  U.S.
;      ZIP:  48334
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/952,014C
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kohn, Kenneth I.
;      REGISTRATION NUMBER:  30,995
;      REFERENCE/DOCKET NUMBER:  2290.00028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  810-539-5050
;      TELEFAX:  810-539-5055
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  500 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-952-014C-18

  Query Match              0.5%;  Score 17;  DB 4;  Length 500;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      157 GCAACAACAGCGAAATC 173
            |||||||||||||||||
Db       43 GCAACAACAGCGAAATC 59


RESULT 20
US-08-998-416-826
; Sequence 826, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
```

```
;       CITY:  Research Triangle Park
;       STATE:  No. 6239264th Carolina
;       COUNTRY:  USA
;       ZIP:  27709
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/998,416
;       FILING DATE:  24-DEC-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  CH 0016/97
;       FILING DATE:  31-DEC-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meigs, J. Timothy
;       REGISTRATION NUMBER:  38,241
;       REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  919-541-8587
;       TELEFAX:  919-541-8689
;   INFORMATION FOR SEQ ID NO:  826:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  671 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     ORIGINAL SOURCE:
;       ORGANISM:  PAG1524UP
US-08-998-416-826

  Query Match              0.5%;  Score 17;  DB 4;  Length 671;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2465 GCAAAATGCGCGGCAGT 2481
           |||||||||||||||||
Db     589 GCAAAATGCGCGGCAGT 605


RESULT 21
PCT-US95-04801-4
; Sequence 4, Application PC/TUS9504801
;  GENERAL INFORMATION:
;     APPLICANT:  Martin, Juan F.
;     APPLICANT:  Coque, Juan R.
;     APPLICANT:  Enguita, Francisco J.
;     APPLICANT:  Fuente, Juan L.
;     APPLICANT:  Llarena, Francisco J.
;     APPLICANT:  Liras, Paloma
;     TITLE OF INVENTION:  DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
;     TITLE OF INVENTION:  LATE GENES
;     NUMBER OF SEQUENCES:  8
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  John W. Wallen III
;       STREET:  P.O. Box 2000
;       CITY:  Rahway
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07065
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/04801
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Wallen III, John W.
;       REGISTRATION NUMBER:  35,403
;       REFERENCE/DOCKET NUMBER:  19179
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 594-3905
;       TELEFAX:  (908) 594-4720
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  800 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
PCT-US95-04801-4

  Query Match              0.5%;  Score 17;  DB 5;  Length 800;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2129 ACCTGATGGTCGAACTG 2145
           |||||||||||||||||
Db     307 ACCTGATGGTCGAACTG 323


RESULT 22
US-08-906-156A-13
; Sequence 13, Application US/08906156A
; Patent No. 6287854
;  GENERAL INFORMATION:
;     APPLICANT:  SPURR, NIGEL K
;     APPLICANT:  GRAY, IAN C
;     APPLICANT:  STEWART, LORNA M
;     TITLE OF INVENTION:  DIAGNOSIS OF SUSCEPTIBILITY TO CANCER
;     TITLE OF INVENTION:  AND TREATMENT THEREOF
;     NUMBER OF SEQUENCES:  94
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  NIXON & VANDERHYE P.C.
;       STREET:  1100 NORTH GLEBE ROAD
;       CITY:  ARLINGTON
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22201
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/906,156A
;       FILING DATE:  05-AUG-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/042,655
;       FILING DATE:  02-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/033,147
;       FILING DATE:  13-DEC-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/005,840
;       FILING DATE:  23-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/96GB/02588
;       FILING DATE:  22-OCT-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  SADOFF, B.J.
;       REGISTRATION NUMBER:  36,663
;       REFERENCE/DOCKET NUMBER:  1090-14
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-816-4000
;       TELEFAX:  703-816-4100
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  871 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Part of gene corresponding to IMAGE 264611
US-08-906-156A-13

  Query Match             0.5%;  Score 17;  DB 4;  Length 871;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         88 GGCGGCGGCGGCGGCAC 104
              |||||||||||||||||
Db        112 GGCGGCGGCGGCGGCAC 128


RESULT 23
US-08-906-156A-58
; Sequence 58, Application US/08906156A
; Patent No. 6287854
;  GENERAL INFORMATION:
;    APPLICANT:  SPURR, NIGEL K
;    APPLICANT:  GRAY, IAN C
;    APPLICANT:  STEWART, LORNA M
;    TITLE OF INVENTION:  DIAGNOSIS OF SUSCEPTIBILITY TO CANCER
;    TITLE OF INVENTION:  AND TREATMENT THEREOF
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/906,156A
;      FILING DATE:  05-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/042,655
;      FILING DATE:  02-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,147
;      FILING DATE:  13-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,840
;      FILING DATE:  23-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/96GB/02588
;      FILING DATE:  22-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SADOFF, B.J.
;      REGISTRATION NUMBER:  36,663
;      REFERENCE/DOCKET NUMBER:  1090-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  871 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  human
;    IMMEDIATE SOURCE:
;      CLONE:  part of IMAGE clone 264611
US-08-906-156A-58

  Query Match             0.5%;  Score 17;  DB 4;  Length 871;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         88 GGCGGCGGCGGCGGCAC 104
              |||||||||||||||||
Db        112 GGCGGCGGCGGCGGCAC 128


RESULT 24
US-09-267-031-5
; Sequence 5, Application US/09267031
; Patent No. 6137031
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Yuelin
;  APPLICANT: Kinkema, Mark
;  APPLICANT: Dong, Xinnian
;  APPLICANT: Ronald, Pamela
;  APPLICANT: Chern, MawShenq
;  APPLICANT: The Regents of the University of California
;  APPLICANT: Duke University
;  TITLE OF INVENTION: DNA Binding Proteins That Interact With NPR1
;  FILE REFERENCE: 023070-092500US
;  CURRENT APPLICATION NUMBER: US/09/267,031
;  CURRENT FILING DATE: 1999-03-11
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 925
;   TYPE: DNA
;   ORGANISM: Oryza sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(925)
;   OTHER INFORMATION: rice bZIP gene MN38
US-09-267-031-5

  Query Match             0.5%;  Score 17;  DB 3;  Length 925;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        545 TTAAAGCTTCTTTCGAC 561
              |||||||||||||||||
Db        527 TTAAAGCTTCTTTCGAC 543


RESULT 25
US-08-770-565-1/c
; Sequence 1, Application US/08770565
; Patent No. 5846723
;  GENERAL INFORMATION:
;    APPLICANT:  Kim, Nam Woo
;    APPLICANT:  Wu, Fred
;    APPLICANT:  Kealey, James T.
;    APPLICANT:  Pruzan, Ronald
;    APPLICANT:  Weinrich, Scott L.
;    TITLE OF INVENTION:  Methods for Detecting the RNA Component of
;    TITLE OF INVENTION:  Telomerase
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/770,565
;      FILING DATE:  20-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Storella, John R.
;      REGISTRATION NUMBER:  32,944
;      REFERENCE/DOCKET NUMBER:  015389-002300US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-770-565-1

  Query Match             0.5%;  Score 17;  DB 2;  Length 981;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2188 GCGGCCGCCGACCGCAC 2204
          |||||||||||||||||
Db    231 GCGGCCGCCGACCGCAC 215


RESULT 26
US-08-710-249-5/c
; Sequence 5, Application US/08710249
; Patent No. 5858777
;  GENERAL INFORMATION:
;    APPLICANT:  Villeponteau, Bryant
;    APPLICANT:  Feng, Junli
;    APPLICANT:  Andrews, William H.
;    APPLICANT:  Adams, Robert R.
;    TITLE OF INVENTION:  Methods and Reagents for Regulating
;    TITLE OF INVENTION:  Telomere Length and Telomerase Activity
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/710,249
;      FILING DATE:  13-SEP-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/583,808
;      FILING DATE:  05-JAN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/003,492
;      FILING DATE:  08-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Storella, John R.
;      REGISTRATION NUMBER:  32,944
;      REFERENCE/DOCKET NUMBER:  015389-001220US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_RNA
;      LOCATION:  267..715
;      OTHER INFORMATION:  /product= "hTR"
;      OTHER INFORMATION:  /note= "hTR transcript serves as
;      OTHER INFORMATION:  template in the telomerase
;      OTHER INFORMATION:  ribonucleoprotein"
US-08-710-249-5

  Query Match             0.5%;  Score 17;  DB 2;  Length 981;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2188 GCGGCCGCCGACCGCAC 2204
          |||||||||||||||||
Db    231 GCGGCCGCCGACCGCAC 215


RESULT 27
US-08-833-377-1/c
; Sequence 1, Application US/08833377
; Patent No. 5968506
;  GENERAL INFORMATION:
;    APPLICANT:  Weinrich, Scott L.
;    APPLICANT:  Atkinson III, Edward M.
;    APPLICANT:  Lichtsteiner, Serge P.
;    APPLICANT:  Vasserot, Alain P.
;    APPLICANT:  Pruzan, Ronald A.
;    APPLICANT:  Kealey, James T.
;    TITLE OF INVENTION:  Purified Telomerase
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/833,377
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/510,736
;      FILING DATE:  04-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Storella, John R.
;      REGISTRATION NUMBER:  32,944
;      REFERENCE/DOCKET NUMBER:  015389-001110US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
```

```
;     NAME/KEY:  -
;     LOCATION:  1..981
;     OTHER INFORMATION:  /note= "PstI fragment of the 2.4 kb
;     OTHER INFORMATION:  SauIIIA1-HindIII fragment of clone 28-1"
US-08-833-377-1

  Query Match             0.5%;  Score 17;  DB 2;  Length 981;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2188 GCGGCCGCCGACCGCAC 2204
           |||||||||||||||||
Db     231 GCGGCCGCCGACCGCAC 215


RESULT 28
US-08-714-482-2/c
; Sequence 2, Application US/08714482
; Patent No. 5972605
;  GENERAL INFORMATION:
;    APPLICANT:  Villeponteau, Bryant
;    APPLICANT:  Harley, Calvin
;    TITLE OF INVENTION:  Assays for Regulators of Mammalian
;    TITLE OF INVENTION:  Telomerase Expression
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/714,482
;      FILING DATE:  16-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/521,634
;      FILING DATE:  31-AUG-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,115
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,802
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/330,123
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/272,102
;      FILING DATE:  07-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Storella, John R.
;      REGISTRATION NUMBER:  32,944
;      REFERENCE/DOCKET NUMBER:  015389-00860US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..981
;      OTHER INFORMATION:  /note= "PstI fragment containing hTR
;      OTHER INFORMATION:  sequence"
US-08-714-482-2

  Query Match             0.5%;  Score 17;  DB 2;  Length 981;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2188 GCGGCCGCCGACCGCAC 2204
           |||||||||||||||||
Db     231 GCGGCCGCCGACCGCAC 215


RESULT 29
US-08-838-545-22/c
; Sequence 22, Application US/08838545
; Patent No. 6046307
;  GENERAL INFORMATION:
;    APPLICANT:  Shay, Jerry W.
;    APPLICANT:  Wright, Woodring E.
;    APPLICANT:  Piatyszek, Mieczyslaw A.
;    APPLICANT:  Corey, David R.
;    APPLICANT:  No. 6046307ton, James C.
;    TITLE OF INVENTION:  Modulation of Mammalian Telomerase by
;    TITLE OF INVENTION:  Peptide Nucleic Acids
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/838,545
;      FILING DATE:  09-APR-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,019
;      FILING DATE:  09-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Storella, John R.
;      REGISTRATION NUMBER:  32,944
;      REFERENCE/DOCKET NUMBER:  015389-001610US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_RNA
;      LOCATION:  266..716
;      OTHER INFORMATION:  /product= "hTR"
;      OTHER INFORMATION:  /note= "RNA component of human telomerase (hTR)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..981
;      OTHER INFORMATION:  /note= "PstI fragment of
;      OTHER INFORMATION:  SauIIIA1-HindIII fragment of clone 28-1"
US-08-838-545-22

  Query Match             0.5%;  Score 17;  DB 3;  Length 981;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
```

```
  Matches    17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2188 GCGGCCGCCGACCGCAC 2204
           ||||||||||||||||
Db     231 GCGGCCGCCGACCGCAC 215

RESULT 30
US-09-349-532-22/c
; Sequence 22, Application US/09349532
; Patent No. 6294650
;  GENERAL INFORMATION:
;    APPLICANT:  Shay, Jerry W.
;    APPLICANT:  Wright, Woodring E.
;    APPLICANT:  Piatyszek, Mieczyslaw A.
;    APPLICANT:  Corey, David R.
;    APPLICANT:  No. 6294650ton, James C.
;    TITLE OF INVENTION:  Modulation of Mammalian Telomerase by
;    TITLE OF INVENTION:  Peptide Nucleic Acids
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/349,532
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/838,545
;      FILING DATE:  09-APR-1997
;      APPLICATION NUMBER:  US 08/630,019
;      FILING DATE:  09-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Storella, John R.
;      REGISTRATION NUMBER:  32,944
;      REFERENCE/DOCKET NUMBER:  015389-001610US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_RNA
;      LOCATION:  266..716
;      OTHER INFORMATION:  /product= "hTR"
;      OTHER INFORMATION:  /note= "RNA component of human telomerase (hTR)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..981
;      OTHER INFORMATION:  /note= "PstI fragment of
;      OTHER INFORMATION:  SauIIIA1-HindIII fragment of clone 28-1"
US-09-349-532-22

  Query Match             0.5%;  Score 17;  DB 4;  Length 981;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2188 GCGGCCGCCGACCGCAC 2204
           ||||||||||||||||
Db     231 GCGGCCGCCGACCGCAC 215

RESULT 31
US-09-220-157A-5/c
; Sequence 5, Application US/09220157A
; Patent No. 6300110
;  GENERAL INFORMATION:
;    APPLICANT:  Villeponteau, Bryant
;    APPLICANT:  Feng, Junli
;    APPLICANT:  Andrews, William H.
;    APPLICANT:  Adams, Robert R.
;    TITLE OF INVENTION:  Methods and Reagents for Regulating
;    TITLE OF INVENTION:  Telomere Length and Telomerase Activity
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/220,157A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/710,249
;      FILING DATE:  13-SEP-1996
;      APPLICATION NUMBER:  US 08/583,808
;      FILING DATE:  05-JAN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/003,492
;      FILING DATE:  08-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Storella, John R.
;      REGISTRATION NUMBER:  32,944
;      REFERENCE/DOCKET NUMBER:  015389-001220US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_RNA
;      LOCATION:  267..715
;      OTHER INFORMATION:  /product= "hTR"
;      OTHER INFORMATION:  /note= "hTR transcript serves as
;      OTHER INFORMATION:  template in the telomerase
;      OTHER INFORMATION:  ribonucleoprotein"
US-09-220-157A-5

  Query Match             0.5%;  Score 17;  DB 4;  Length 981;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2188 GCGGCCGCCGACCGCAC 2204
           ||||||||||||||||
Db     231 GCGGCCGCCGACCGCAC 215

RESULT 32
US-07-601-094-30/c
```

```
; Sequence 30, Application US/07601094
; Patent No. 5215892
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Hirano, Toshio
;    APPLICANT:  Akira, Shizuo
;    APPLICANT:  Isshiki, Hiroshi
;    APPLICANT:  Tanabe, Osamu
;    APPLICANT:  Kinoshita, Shigemi
;    APPLICANT:  Shimamoto, Takuya
;    TITLE OF INVENTION:  C/EBP2 Gene and Recombinant
;    TITLE OF INVENTION:  C/EBP2
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak &
;      ADDRESSEE:  Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/601,094
;      FILING DATE:  19901022
;      CLASSIFICATION:  435
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1035 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1035
;      OTHER INFORMATION:
US-07-601-094-30

  Query Match             0.5%;  Score 17;  DB 1;  Length 1035;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db      57 AGGCGGCGGCGGCGGCA 41


RESULT 33
US-08-012-735-30/c
; Sequence 30, Application US/08012735
; Patent No. 5360894
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Hirano, Toshio
;    APPLICANT:  Akira, Shizuo
;    APPLICANT:  Isshiki, Hiroshi
;    APPLICANT:  Tanabe, Osamu
;    APPLICANT:  Kinoshita, Shigemi
;    APPLICANT:  Shimamoto, Takuya
;    TITLE OF INVENTION:  C/EBP2 Gene and Recombinant
;    TITLE OF INVENTION:  C/EBP2
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak &
;      ADDRESSEE:  Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/012,735
;      FILING DATE:  19930203
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/601,094
;      FILING DATE:  22 OCT 1990
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1035 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1035
;      OTHER INFORMATION:
US-08-012-735-30

  Query Match             0.5%;  Score 17;  DB 1;  Length 1035;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db      57 AGGCGGCGGCGGCGGCA 41


RESULT 34
US-08-791-115B-8
; Sequence 8, Application US/08791115B
; Patent No. 6262242
;  GENERAL INFORMATION:
;    APPLICANT:  Steck, Peter
;    APPLICANT:  Pershouse, Mark A.
;    APPLICANT:  Jasser, Samar
;    APPLICANT:  Yung, W.K. Alfred
;    APPLICANT:  Tavtigian, Sean V.
;    TITLE OF INVENTION:  A TUMOR SUPPRESSOR DESIGNATED TS10Q23.3
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, P.C.
;      STREET:  555 Thirteenth Street, N.W., Suite 701-E
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  22204
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/791,115B
;      FILING DATE:  30-JAN-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  38,957
;      REFERENCE/DOCKET NUMBER:  2318-134.A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-683-6040
;      TELEFAX:  202-683-7031
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1257 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-791-115B-8

  Query Match             0.5%;  Score 17;  DB 4;  Length 1257;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      88 GGCGGCGGCGGCGGCAC 104
           |||||||||||||||||
Db     490 GGCGGCGGCGGCGGCAC 506


RESULT 35
US-08-224-482-5/c
; Sequence 5, Application US/08224482
; Patent No. 5837692
;  GENERAL INFORMATION:
;    APPLICANT:  Mercola, Dan
;    APPLICANT:  Adamson, Eileen D.
;    TITLE OF INVENTION:  Inhibition of the Mitogenic Activity of
;    TITLE OF INVENTION:  PDGF by Mammalian EGR
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/224,482
;      FILING DATE:  07-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ME 9913
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  57..1422
US-08-224-482-5

  Query Match             0.5%;  Score 17;  DB 2;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      86 TAGGCGGCGGCGGCGGC 102
           |||||||||||||||||
Db     982 TAGGCGGCGGCGGCGGC 966


RESULT 36
US-08-906-156A-11
; Sequence 11, Application US/08906156A
; Patent No. 6287854
;  GENERAL INFORMATION:
;    APPLICANT:  SPURR, NIGEL K
;    APPLICANT:  GRAY, IAN C
;    APPLICANT:  STEWART, LORNA M
;    TITLE OF INVENTION:  DIAGNOSIS OF SUSCEPTIBILITY TO CANCER
;    TITLE OF INVENTION:  AND TREATMENT THEREOF
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/906,156A
;      FILING DATE:  05-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/042,655
;      FILING DATE:  02-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,147
;      FILING DATE:  13-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,840
;      FILING DATE:  23-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/96GB/02588
;      FILING DATE:  22-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SADOFF, B.J.
;      REGISTRATION NUMBER:  36,663
;      REFERENCE/DOCKET NUMBER:  1090-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1773 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Partial cDNA of gene corresponding to IMAGE 264611
US-08-906-156A-11

  Query Match             0.5%;  Score 17;  DB 4;  Length 1773;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      88 GGCGGCGGCGGCGGCAC 104
           |||||||||||||||||
Db     112 GGCGGCGGCGGCGGCAC 128


RESULT 37
US-09-310-363C-7/c
; Sequence 7, Application US/09310363C
```

```
; Patent No. 6388169
; GENERAL INFORMATION:
;  APPLICANT: Mahajan, Pramod B.
;  APPLICANT:  Shi, Jinrui
;  APPLICANT:  McElver, John
;  APPLICANT:  Bowen, Benjamin
;  APPLICANT:  Baszczynski, Christopher
;  TITLE OF INVENTION: RecA cDNAs and Uses Thereof
;  FILE REFERENCE: 0855
;  CURRENT APPLICATION NUMBER: US/09/310,363C
;  CURRENT FILING DATE:  1999-05-12
;  PRIOR APPLICATION NUMBER: 60/099,765
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/096,492
;  PRIOR FILING DATE: 1998-08-14
;  PRIOR APPLICATION NUMBER: 60/088,529
;  PRIOR FILING DATE: 1998-06-08
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 1773
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (97)...(1371)
US-09-310-363C-7

  Query Match             0.5%;  Score 17;  DB 4;  Length 1773;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     87 AGGCGGCGGCGGCGGCA 103
          |||||||||||||||||
Db    154 AGGCGGCGGCGGCGGCA 138


RESULT 38
PCT-US95-02455-1
; Sequence 1, Application PC/TUS9502455
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs et al
;    TITLE OF INVENTION:  GENE FOR MYCOBACTERIAL
;    TITLE OF INVENTION:  DIAMINOPIMELIC ACID
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amster, Rothstein & Ebenstein
;      STREET:  90 Park Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch 1.44 Mb
;      MEDIUM TYPE:  storage diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (ASCII)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/02455
;      FILING DATE:  Not Yet Assigned
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:  none
;      APPLICATION NUMBER:  none
;      FILING DATE:  none
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  George, Kenneth P.
;      REGISTRATION NUMBER:  30,259
;      REFERENCE/DOCKET NUMBER:  96700/358
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1801
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  dapB gene
;    HYPOTHETICAL:  no
;    ANTI-SENSE:
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:  BCG
;      ORGANISM:  BCG
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:
;      CELL LINE:
;      ORGANELLE:
;    IMMEDIATE SOURCE:
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
;    PUBLICATION INFORMATION:  none
;      AUTHORS:
;      TITLE:
;      JOURNAL:
;      VOLUME:
;      PAGES:
;      DATE:
;      DOCUMENT NUMBER:
;      FILING DATE:
;      PUBLICATION DATE:
;      RELEVANT RESIDUES IN SEQ ID NO:
PCT-US95-02455-1

  Query Match             0.5%;  Score 17;  DB 5;  Length 1801;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     88 GGCGGCGGCGGCGGCAC 104
          |||||||||||||||||
Db    663 GGCGGCGGCGGCGGCAC 679


RESULT 39
US-09-593-711A-3/c
; Sequence 3, Application US/09593711A
; Patent No. 6271030
; GENERAL INFORMATION:
;  APPLICANT:    Brett P. Monia
;  APPLICANT:  Madeline M. Butler
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION:    ANTISENSE MODULATION OF C/EBP BETA EXPRESSION
;  FILE REFERENCE:    RTS-0118
;  CURRENT APPLICATION NUMBER: US/09/593,711A
;  CURRENT FILING DATE:  2000-06-14
;  NUMBER OF SEQ ID NOS:     244
; SEQ ID NO 3
;   LENGTH: 1910
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION:
;   NAME/KEY: unsure
;   LOCATION: 1415
;   OTHER INFORMATION: unknown
```

```
;   NAME/KEY: unsure
;   LOCATION: 1421
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 1422
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 1423
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 1424
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 1458
;   OTHER INFORMATION: unknown
;   NAME/KEY: CDS
;   LOCATION: (299)...(1336)
US-09-593-711A-3

  Query Match              0.5%;  Score 17;  DB 4;  Length 1910;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db     355 AGGCGGCGGCGGCGGCA 339


RESULT 40
US-07-601-094-1/c
; Sequence 1, Application US/07601094
; Patent No. 5215892
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Hirano, Toshio
;    APPLICANT:  Akira, Shizuo
;    APPLICANT:  Isshiki, Hiroshi
;    APPLICANT:  Tanabe, Osamu
;    APPLICANT:  Kinoshita, Shigemi
;    APPLICANT:  Shimamoto, Takuya
;    TITLE OF INVENTION:  C/EBP2 Gene and Recombinant
;    TITLE OF INVENTION:  C/EBP2
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak &
;      ADDRESSEE:  Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/601,094
;      FILING DATE:  19901022
;      CLASSIFICATION:  435
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1914 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  281..1316
;      OTHER INFORMATION:
US-07-601-094-1

  Query Match              0.5%;  Score 17;  DB 1;  Length 1914;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db     337 AGGCGGCGGCGGCGGCA 321


Search completed: January 27, 2003, 23:07:26
Job time : 163 secs
```

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 23:07:29 ; Search time 104 Seconds
(without alignments)
13840.812 Million cell updates/sec

Title: US-09-830-433A-7
Perfect score: 3204
Sequence: 1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 396772 seqs, 224632407 residues

Word size : 0

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	177	5.5	286	9	US-09-928-457-95	Sequence 95, Appl
	2	19	0.6	284	10	US-09-867-701-2325	Sequence 2325, Ap
	3	19	0.6	1267	10	US-09-925-299-125	Sequence 125, App
c	4	19	0.6	1551	10	US-09-880-107-2355	Sequence 2355, Ap
	5	19	0.6	1987	10	US-09-925-301-306	Sequence 306, App
	6	18	0.6	213	10	US-09-974-300-4730	Sequence 4730, Ap
	7	18	0.6	276	10	US-09-878-574-15701	Sequence 15701, A
c	8	18	0.6	495	10	US-09-728-445-822	Sequence 822, App
c	9	18	0.6	1044	10	US-09-974-300-2315	Sequence 2315, Ap
	10	18	0.6	1169	9	US-10-098-841-169	Sequence 169, App
	11	18	0.6	2321	9	US-10-071-766-34	Sequence 34, Appl
	12	18	0.6	4446	10	US-09-815-242-7983	Sequence 7983, Ap
c	13	18	0.6	29729	10	US-09-070-927A-238	Sequence 238, App
	14	17	0.5	102	10	US-09-789-836-22	Sequence 22, Appl
c	15	17	0.5	105	10	US-09-789-836-21	Sequence 21, Appl
c	16	17	0.5	106	10	US-09-789-836-33	Sequence 33, Appl
	17	17	0.5	122	10	US-09-783-590-5143	Sequence 5143, Ap
c	18	17	0.5	126	10	US-09-974-300-4131	Sequence 4131, Ap
	19	17	0.5	224	10	US-09-878-574-8654	Sequence 8654, Ap
	20	17	0.5	335	10	US-09-880-107-3632	Sequence 3632, Ap
c	21	17	0.5	336	10	US-09-974-300-7591	Sequence 7591, Ap
c	22	17	0.5	381	10	US-09-983-965-4734	Sequence 4734, Ap
c	23	17	0.5	388	9	US-10-040-739-1122	Sequence 1122, Ap
	24	17	0.5	415	10	US-09-983-965-5116	Sequence 5116, Ap
c	25	17	0.5	440	9	US-09-854-133-339	Sequence 339, App
c	26	17	0.5	440	10	US-09-738-973-339	Sequence 339, App
	27	17	0.5	441	10	US-09-815-242-7839	Sequence 7839, Ap
c	28	17	0.5	480	10	US-09-864-761-204	Sequence 204, App
c	29	17	0.5	484	9	US-10-040-739-17	Sequence 17, Appl
c	30	17	0.5	490	10	US-09-881-752A-225	Sequence 225, App
c	31	17	0.5	532	10	US-09-864-761-8567	Sequence 8567, Ap
c	32	17	0.5	556	9	US-10-108-915-31	Sequence 31, Appl
c	33	17	0.5	578	10	US-09-864-761-12180	Sequence 12180, A
	34	17	0.5	747	9	US-09-738-626-2896	Sequence 2896, Ap
c	35	17	0.5	872	10	US-09-764-864-326	Sequence 326, App
	36	17	0.5	915	9	US-09-938-842A-2261	Sequence 2261, Ap
	37	17	0.5	960	10	US-09-974-300-5167	Sequence 5167, Ap
c	38	17	0.5	1038	10	US-09-789-836-2	Sequence 2, Appli
c	39	17	0.5	1408	10	US-09-822-830A-366	Sequence 366, App
c	40	17	0.5	1428	9	US-10-108-915-33	Sequence 33, Appl
	41	17	0.5	1545	10	US-09-815-242-7675	Sequence 7675, Ap
	42	17	0.5	1610	10	US-09-864-864-304	Sequence 304, App
	43	17	0.5	1706	12	US-10-033-109-3	Sequence 3, Appli
	44	17	0.5	1768	10	US-09-833-381-523	Sequence 523, App
c	45	17	0.5	1816	9	US-09-974-298-18	Sequence 18, Appl
c	46	17	0.5	1910	10	US-09-789-836-1	Sequence 1, Appli
	47	17	0.5	1967	10	US-09-880-107-2417	Sequence 2417, Ap
c	48	17	0.5	2016	10	US-09-764-864-744	Sequence 744, App
c	49	17	0.5	2037	9	US-09-712-363-12	Sequence 12, Appl
	50	17	0.5	2139	12	US-10-028-780-1	Sequence 1, Appli
c	51	17	0.5	2426	10	US-09-057-351-3	Sequence 3, Appli
	52	17	0.5	2471	10	US-09-915-060-3	Sequence 3, Appli
	53	17	0.5	2617	10	US-09-925-302-190	Sequence 190, App
	54	17	0.5	2796	10	US-09-871-388-3	Sequence 3, Appli
c	55	17	0.5	2849	10	US-09-880-107-2110	Sequence 2110, Ap
c	56	17	0.5	2887	9	US-09-964-899-34	Sequence 34, Appl
c	57	17	0.5	2910	10	US-09-833-790-418	Sequence 418, App
c	58	17	0.5	2910	10	US-09-880-107-3784	Sequence 3784, Ap
	59	17	0.5	2930	10	US-09-960-253-156	Sequence 156, App
c	60	17	0.5	2971	12	US-10-002-600-42	Sequence 42, Appl
	61	17	0.5	3044	10	US-09-880-107-3718	Sequence 3718, Ap
	62	17	0.5	3047	10	US-09-864-864-329	Sequence 329, App
	63	17	0.5	3115	10	US-09-925-299-123	Sequence 123, App
	64	17	0.5	3402	10	US-09-949-192-18	Sequence 18, Appl
c	65	17	0.5	3531	9	US-09-992-598-32	Sequence 32, Appl
c	66	17	0.5	3531	9	US-09-989-293A-32	Sequence 32, Appl
c	67	17	0.5	3531	9	US-09-989-735-32	Sequence 32, Appl
c	68	17	0.5	3531	9	US-09-990-444-32	Sequence 32, Appl
c	69	17	0.5	3531	9	US-09-989-730-32	Sequence 32, Appl
c	70	17	0.5	3531	9	US-09-990-436-32	Sequence 32, Appl
c	71	17	0.5	3531	9	US-09-991-181-32	Sequence 32, Appl
c	72	17	0.5	3531	9	US-09-993-687-32	Sequence 32, Appl
c	73	17	0.5	3531	9	US-09-989-734-32	Sequence 32, Appl
c	74	17	0.5	3531	9	US-09-997-653-32	Sequence 32, Appl
c	75	17	0.5	3531	10	US-09-989-722-32	Sequence 32, Appl
c	76	17	0.5	3531	10	US-09-989-723-32	Sequence 32, Appl
c	77	17	0.5	3531	10	US-09-989-279-32	Sequence 32, Appl
c	78	17	0.5	3531	10	US-09-989-727-32	Sequence 32, Appl
c	79	17	0.5	3531	10	US-09-989-731-32	Sequence 32, Appl
c	80	17	0.5	3531	10	US-09-989-732-32	Sequence 32, Appl
c	81	17	0.5	3531	10	US-09-991-073-32	Sequence 32, Appl
c	82	17	0.5	3531	10	US-09-990-442-32	Sequence 32, Appl
c	83	17	0.5	3531	10	US-09-991-163-32	Sequence 32, Appl
c	84	17	0.5	3531	10	US-09-993-604-32	Sequence 32, Appl
c	85	17	0.5	3531	10	US-09-990-456-32	Sequence 32, Appl
c	86	17	0.5	3531	10	US-09-989-721-32	Sequence 32, Appl
	87	17	0.5	3622	12	US-10-002-600-5	Sequence 5, Appli
	88	17	0.5	3656	9	US-10-167-264-1	Sequence 1, Appli
c	89	17	0.5	3969	10	US-09-880-107-2213	Sequence 2213, Ap
	90	17	0.5	4138	10	US-09-920-804-3	Sequence 3, Appli
c	91	17	0.5	4534	9	US-10-114-091-1	Sequence 1, Appli
	92	17	0.5	4641	10	US-09-925-300-686	Sequence 686, App

	93	17	0.5	4751	10	US-09-764-864-1757	Sequence 1757, Ap
c	94	17	0.5	4957	10	US-09-880-107-2154	Sequence 2154, Ap
	95	17	0.5	5385	10	US-09-920-804-1	Sequence 1, Appli
	96	17	0.5	5565	9	US-10-029-413A-1	Sequence 1, Appli
	97	17	0.5	5735	10	US-09-734-674-1	Sequence 1, Appli
	98	17	0.5	6160	9	US-10-029-413A-11	Sequence 11, Appl
	99	17	0.5	6160	9	US-10-029-413A-13	Sequence 13, Appl
	100	17	0.5	6615	9	US-10-029-413A-9	Sequence 9, Appli
	101	17	0.5	8147	12	US-10-109-886-9	Sequence 9, Appli
	102	17	0.5	8711	10	US-09-764-864-1783	Sequence 1783, Ap
	103	17	0.5	9592	9	US-10-167-264-3	Sequence 3, Appli
	104	17	0.5	9990	10	US-09-767-088A-2	Sequence 2, Appli
	105	17	0.5	9990	10	US-09-767-088A-15	Sequence 15, Appl
	106	17	0.5	13048	10	US-09-880-107-2364	Sequence 2364, Ap
	107	17	0.5	14093	10	US-09-764-847-1744	Sequence 1744, Ap
c	108	17	0.5	18871	10	US-09-764-847-1317	Sequence 1317, Ap
	109	17	0.5	26657	10	US-09-810-673A-3	Sequence 3, Appli
	110	17	0.5	35937	10	US-09-782-378A-3	Sequence 3, Appli
	111	17	0.5	202001	10	US-09-734-674-3	Sequence 3, Appli
	112	17	0.5	368004	10	US-09-949-654-3	Sequence 3, Appli
c	113	17	0.5	3309400	9	US-09-738-626-1	Sequence 1, Appli
	114	16	0.5	29	8	US-08-834-666A-47	Sequence 47, Appl
	115	16	0.5	30	8	US-08-834-666A-44	Sequence 44, Appl
	116	16	0.5	99	10	US-09-864-761-31942	Sequence 31942, A
	117	16	0.5	142	10	US-09-770-696-549	Sequence 549, App
c	118	16	0.5	146	10	US-09-864-761-32039	Sequence 32039, A
c	119	16	0.5	207	10	US-09-878-574-13143	Sequence 13143, A
	120	16	0.5	209	10	US-09-770-696-101	Sequence 101, App
	121	16	0.5	210	10	US-09-864-761-22135	Sequence 22135, A
c	122	16	0.5	230	10	US-09-864-761-23975	Sequence 23975, A
	123	16	0.5	234	10	US-09-960-352-11055	Sequence 11055, A
c	124	16	0.5	243	10	US-09-764-878-34	Sequence 34, Appl
	125	16	0.5	256	9	US-09-933-797-425	Sequence 425, App
c	126	16	0.5	256	9	US-10-040-739-812	Sequence 812, App
	127	16	0.5	257	10	US-09-878-574-12851	Sequence 12851, A
	128	16	0.5	257	10	US-09-878-574-15531	Sequence 15531, A
c	129	16	0.5	263	10	US-09-878-574-15376	Sequence 15376, A
c	130	16	0.5	264	10	US-09-878-574-6336	Sequence 6336, Ap
c	131	16	0.5	264	10	US-09-878-574-6353	Sequence 6353, Ap
c	132	16	0.5	264	10	US-09-878-574-14016	Sequence 14016, A
c	133	16	0.5	264	10	US-09-878-574-15044	Sequence 15044, A
c	134	16	0.5	265	10	US-09-878-574-9190	Sequence 9190, Ap
c	135	16	0.5	267	10	US-09-764-860-795	Sequence 795, App
c	136	16	0.5	267	10	US-09-878-574-15077	Sequence 15077, A
c	137	16	0.5	268	10	US-09-878-574-6774	Sequence 6774, Ap
c	138	16	0.5	268	10	US-09-878-574-11810	Sequence 11810, A
c	139	16	0.5	269	10	US-09-878-574-12677	Sequence 12677, A
c	140	16	0.5	271	10	US-09-764-860-98	Sequence 98, Appl
	141	16	0.5	271	10	US-09-878-574-8045	Sequence 8045, Ap
c	142	16	0.5	273	10	US-09-878-574-11691	Sequence 11691, A
	143	16	0.5	273	10	US-09-878-574-11702	Sequence 11702, A
	144	16	0.5	274	10	US-09-294-093B-732	Sequence 732, App
	145	16	0.5	274	10	US-09-923-876-4950	Sequence 4950, Ap
	146	16	0.5	275	10	US-09-878-574-8772	Sequence 8772, Ap
	147	16	0.5	276	10	US-09-294-093B-604	Sequence 604, App
c	148	16	0.5	279	10	US-09-878-574-5482	Sequence 5482, Ap
	149	16	0.5	283	10	US-09-923-876-4855	Sequence 4855, Ap
	150	16	0.5	283	10	US-09-917-800A-688	Sequence 688, App
c	151	16	0.5	284	9	US-10-046-935-682	Sequence 682, App
c	152	16	0.5	284	9	US-09-878-178-682	Sequence 682, App
c	153	16	0.5	284	10	US-09-880-107-1966	Sequence 1966, Ap
c	154	16	0.5	286	10	US-09-878-574-7994	Sequence 7994, Ap
c	155	16	0.5	298	10	US-09-983-965-5793	Sequence 5793, Ap
	156	16	0.5	301	10	US-09-974-300-6395	Sequence 6395, Ap
c	157	16	0.5	304	10	US-09-294-093B-532	Sequence 532, App
	158	16	0.5	310	10	US-09-867-701-4186	Sequence 4186, Ap
c	159	16	0.5	314	10	US-09-864-761-32286	Sequence 32286, A
	160	16	0.5	314	10	US-09-783-590-9850	Sequence 9850, Ap
c	161	16	0.5	315	10	US-09-925-299-280	Sequence 280, App
c	162	16	0.5	330	9	US-09-933-797-11	Sequence 11, Appl
c	163	16	0.5	330	9	US-09-933-797-12	Sequence 12, Appl
	164	16	0.5	336	10	US-09-983-965-3727	Sequence 3727, Ap
c	165	16	0.5	339	10	US-09-682-706-1	Sequence 1, Appli
c	166	16	0.5	345	10	US-09-894-882-76	Sequence 76, Appl
	167	16	0.5	351	10	US-09-815-343-389	Sequence 389, App
	168	16	0.5	352	10	US-09-960-352-6025	Sequence 6025, Ap
	169	16	0.5	355	9	US-10-025-380-1026	Sequence 1026, Ap
	170	16	0.5	355	10	US-09-922-217-1026	Sequence 1026, Ap
	171	16	0.5	355	10	US-09-833-263-1026	Sequence 1026, Ap
c	172	16	0.5	363	10	US-09-960-352-2021	Sequence 2021, Ap
c	173	16	0.5	367	10	US-09-864-761-17375	Sequence 17375, A
c	174	16	0.5	370	10	US-09-960-352-12015	Sequence 12015, A
	175	16	0.5	372	10	US-09-745-763-80	Sequence 80, Appl
	176	16	0.5	373	10	US-09-864-761-11275	Sequence 11275, A
c	177	16	0.5	378	10	US-09-867-701-4708	Sequence 4708, Ap
	178	16	0.5	387	9	US-10-025-380-727	Sequence 727, App
	179	16	0.5	387	10	US-09-922-217-727	Sequence 727, App
	180	16	0.5	387	10	US-09-833-263-727	Sequence 727, App
	181	16	0.5	395	10	US-09-878-574-2838	Sequence 2838, Ap
c	182	16	0.5	396	9	US-09-970-966-7	Sequence 7, Appli
c	183	16	0.5	396	9	US-09-970-966-63	Sequence 63, Appl
c	184	16	0.5	396	9	US-09-970-966-64	Sequence 64, Appl
c	185	16	0.5	396	10	US-09-825-294-7	Sequence 7, Appli
c	186	16	0.5	396	10	US-09-825-294-63	Sequence 63, Appl
c	187	16	0.5	396	10	US-09-825-294-64	Sequence 64, Appl
c	188	16	0.5	398	10	US-09-878-574-3153	Sequence 3153, Ap
c	189	16	0.5	399	10	US-09-919-473-1	Sequence 1, Appli
	190	16	0.5	401	9	US-09-946-807-847	Sequence 847, App
	191	16	0.5	401	9	US-09-946-807-848	Sequence 848, App
	192	16	0.5	401	10	US-09-795-668-847	Sequence 847, App
	193	16	0.5	401	10	US-09-795-668-848	Sequence 848, App
	194	16	0.5	401	10	US-09-795-686-847	Sequence 847, App
	195	16	0.5	401	10	US-09-795-686-848	Sequence 848, App
	196	16	0.5	402	9	US-09-854-133-418	Sequence 418, App
	197	16	0.5	402	10	US-09-738-973-418	Sequence 418, App
c	198	16	0.5	406	10	US-09-864-761-3997	Sequence 3997, Ap
	199	16	0.5	412	10	US-09-864-761-11431	Sequence 11431, A
	200	16	0.5	412	10	US-09-960-352-13161	Sequence 13161, A
	201	16	0.5	415	10	US-09-878-574-4032	Sequence 4032, Ap
	202	16	0.5	416	9	US-09-854-133-249	Sequence 249, App
	203	16	0.5	416	10	US-09-738-973-249	Sequence 249, App
	204	16	0.5	418	10	US-09-960-352-13187	Sequence 13187, A
c	205	16	0.5	424	10	US-09-815-343-24	Sequence 24, Appl
	206	16	0.5	424	10	US-09-867-701-10122	Sequence 10122, A
	207	16	0.5	426	10	US-09-960-352-12683	Sequence 12683, A
	208	16	0.5	426	10	US-09-960-352-14523	Sequence 14523, A
	209	16	0.5	428	9	US-09-920-455-62	Sequence 62, Appl
	210	16	0.5	429	10	US-09-960-352-6763	Sequence 6763, Ap
	211	16	0.5	430	10	US-09-974-300-3346	Sequence 3346, Ap
	212	16	0.5	440	10	US-09-876-889-328	Sequence 328, App
	213	16	0.5	440	10	US-09-960-352-12927	Sequence 12927, A
	214	16	0.5	441	9	US-09-854-133-304	Sequence 304, App
	215	16	0.5	441	10	US-09-738-973-304	Sequence 304, App
	216	16	0.5	447	10	US-09-864-761-6328	Sequence 6328, Ap
c	217	16	0.5	447	10	US-09-864-761-15527	Sequence 15527, A
	218	16	0.5	447	10	US-09-960-352-5311	Sequence 5311, Ap
	219	16	0.5	451	10	US-09-867-701-2802	Sequence 2802, Ap
	220	16	0.5	456	10	US-09-864-761-1264	Sequence 1264, Ap
c	221	16	0.5	462	10	US-09-974-300-7119	Sequence 7119, Ap
	222	16	0.5	464	10	US-09-864-761-5361	Sequence 5361, Ap
c	223	16	0.5	469	10	US-09-864-761-15780	Sequence 15780, A
c	224	16	0.5	469	10	US-09-864-761-16259	Sequence 16259, A
	225	16	0.5	471	10	US-09-920-300A-1002	Sequence 1002, Ap
c	226	16	0.5	471	10	US-09-974-300-5817	Sequence 5817, Ap
	227	16	0.5	471	12	US-10-033-528-1002	Sequence 1002, Ap
	228	16	0.5	478	9	US-09-854-133-499	Sequence 499, App
	229	16	0.5	478	10	US-09-738-973-499	Sequence 499, App
	230	16	0.5	479	10	US-09-815-343-923	Sequence 923, App
	231	16	0.5	480	10	US-09-864-761-15427	Sequence 15427, A
c	232	16	0.5	491	10	US-09-735-705-307	Sequence 307, App
c	233	16	0.5	491	10	US-09-850-716A-307	Sequence 307, App
c	234	16	0.5	491	10	US-09-897-778-307	Sequence 307, App
	235	16	0.5	496	10	US-09-783-590-3813	Sequence 3813, Ap
	236	16	0.5	503	10	US-09-864-761-7863	Sequence 7863, Ap
	237	16	0.5	509	10	US-09-783-590-3846	Sequence 3846, Ap
	238	16	0.5	511	9	US-09-854-133-359	Sequence 359, App

	239	16	0.5	511	10	US-09-939-980-182	Sequence 182, App
	240	16	0.5	511	10	US-09-738-973-359	Sequence 359, App
c	241	16	0.5	514	9	US-09-933-797-761	Sequence 761, App
	242	16	0.5	521	10	US-09-864-761-8078	Sequence 8078, Ap
c	243	16	0.5	523	10	US-09-864-761-8631	Sequence 8631, Ap
c	244	16	0.5	545	10	US-09-919-580-587	Sequence 587, App
	245	16	0.5	546	10	US-09-998-598-1151	Sequence 1151, Ap
	246	16	0.5	547	10	US-09-867-701-1325	Sequence 1325, Ap
c	247	16	0.5	550	10	US-09-919-580-614	Sequence 614, App
c	248	16	0.5	564	9	US-10-040-739-1008	Sequence 1008, Ap
	249	16	0.5	564	10	US-09-867-701-2010	Sequence 2010, Ap
	250	16	0.5	567	10	US-09-864-761-8352	Sequence 8352, Ap
c	251	16	0.5	567	10	US-09-919-580-605	Sequence 605, App
	252	16	0.5	569	10	US-09-864-761-12956	Sequence 12956, A
	253	16	0.5	577	10	US-09-867-701-2139	Sequence 2139, Ap
c	254	16	0.5	578	10	US-09-864-761-16422	Sequence 16422, A
	255	16	0.5	579	10	US-09-864-761-13007	Sequence 13007, A
	256	16	0.5	588	9	US-10-164-871-1	Sequence 1, Appli
c	257	16	0.5	593	10	US-09-919-580-504	Sequence 504, App
c	258	16	0.5	599	10	US-09-864-761-7244	Sequence 7244, Ap
	259	16	0.5	599	10	US-09-864-761-8950	Sequence 8950, Ap
c	260	16	0.5	602	10	US-09-770-149-955	Sequence 955, App
	261	16	0.5	602	10	US-09-772-134B-34	Sequence 34, Appl
	262	16	0.5	602	10	US-09-772-134B-75	Sequence 75, Appl
c	263	16	0.5	603	9	US-09-764-904-72	Sequence 72, Appl
c	264	16	0.5	603	10	US-09-764-860-600	Sequence 600, App
c	265	16	0.5	607	10	US-09-864-864-242	Sequence 242, App
	266	16	0.5	643	10	US-09-815-343-408	Sequence 408, App
	267	16	0.5	644	10	US-09-772-134B-35	Sequence 35, Appl
c	268	16	0.5	646	10	US-09-867-550-1769	Sequence 1769, Ap
c	269	16	0.5	649	10	US-09-919-580-29	Sequence 29, Appl
c	270	16	0.5	649	10	US-09-772-134B-16	Sequence 16, Appl
c	271	16	0.5	654	10	US-09-813-358-10	Sequence 10, Appl
	272	16	0.5	661	10	US-09-772-134B-86	Sequence 86, Appl
c	273	16	0.5	704	10	US-09-950-688-1	Sequence 1, Appli
c	274	16	0.5	707	9	US-10-098-841-19	Sequence 19, Appl
c	275	16	0.5	728	10	US-09-764-898-78	Sequence 78, Appl
c	276	16	0.5	735	10	US-09-728-446-721	Sequence 721, App
c	277	16	0.5	739	10	US-09-752-639-2	Sequence 2, Appli
c	278	16	0.5	739	10	US-09-984-198-2	Sequence 2, Appli
c	279	16	0.5	740	10	US-09-974-300-4231	Sequence 4231, Ap
c	280	16	0.5	744	9	US-09-938-842A-1922	Sequence 1922, Ap
c	281	16	0.5	747	10	US-09-974-300-2802	Sequence 2802, Ap
	282	16	0.5	768	9	US-09-971-536-14	Sequence 14, Appl
	283	16	0.5	774	9	US-09-738-626-1292	Sequence 1292, Ap
c	284	16	0.5	776	10	US-09-919-580-889	Sequence 889, App
c	285	16	0.5	791	10	US-09-919-580-912	Sequence 912, App
c	286	16	0.5	794	10	US-09-772-134B-18	Sequence 18, Appl
c	287	16	0.5	825	12	US-10-062-254-163	Sequence 163, App
c	288	16	0.5	831	10	US-09-880-503-14	Sequence 14, Appl
c	289	16	0.5	847	9	US-09-890-813-11	Sequence 11, Appl
c	290	16	0.5	860	10	US-09-867-550-1249	Sequence 1249, Ap
c	291	16	0.5	861	10	US-09-772-134B-20	Sequence 20, Appl
	292	16	0.5	885	10	US-09-974-300-5111	Sequence 5111, Ap
c	293	16	0.5	887	7	US-08-841-636A-36	Sequence 36, Appl
c	294	16	0.5	903	10	US-09-919-580-927	Sequence 927, App
	295	16	0.5	906	10	US-09-815-242-7590	Sequence 7590, Ap
c	296	16	0.5	920	10	US-09-765-231A-3	Sequence 3, Appli
c	297	16	0.5	937	12	US-10-062-254-165	Sequence 165, App
c	298	16	0.5	939	10	US-09-974-300-572	Sequence 572, App
c	299	16	0.5	954	9	US-09-738-626-32	Sequence 32, Appl
c	300	16	0.5	954	9	US-09-738-626-183	Sequence 183, App
c	301	16	0.5	957	10	US-09-822-830A-259	Sequence 259, App
c	302	16	0.5	972	10	US-09-880-503-16	Sequence 16, Appl
c	303	16	0.5	972	10	US-09-749-728B-10	Sequence 10, Appl
	304	16	0.5	1002	12	US-10-062-254-203	Sequence 203, App
	305	16	0.5	1006	9	US-09-946-807-104	Sequence 104, App
	306	16	0.5	1006	10	US-09-795-668-104	Sequence 104, App
	307	16	0.5	1006	10	US-09-795-686-104	Sequence 104, App
c	308	16	0.5	1017	9	US-09-738-626-485	Sequence 485, App
	309	16	0.5	1020	9	US-09-712-363-53	Sequence 53, Appl
	310	16	0.5	1032	10	US-09-771-161A-79	Sequence 79, Appl
	311	16	0.5	1035	10	US-09-866-582-5	Sequence 5, Appli
c	312	16	0.5	1045	12	US-10-062-254-161	Sequence 161, App
	313	16	0.5	1071	10	US-09-815-242-7426	Sequence 7426, Ap
c	314	16	0.5	1080	10	US-09-919-580-909	Sequence 909, App
	315	16	0.5	1089	10	US-09-815-242-7255	Sequence 7255, Ap
c	316	16	0.5	1091	10	US-09-822-849A-388	Sequence 388, App
c	317	16	0.5	1113	12	US-10-062-254-205	Sequence 205, App
c	318	16	0.5	1128	10	US-09-764-878-426	Sequence 426, App
c	319	16	0.5	1161	9	US-09-910-186A-25	Sequence 25, Appl
c	320	16	0.5	1161	10	US-09-833-381-2046	Sequence 2046, Ap
	321	16	0.5	1176	10	US-09-815-242-7814	Sequence 7814, Ap
	322	16	0.5	1179	10	US-09-912-020-188	Sequence 188, App
	323	16	0.5	1187	9	US-09-938-842A-5303	Sequence 5303, Ap
	324	16	0.5	1200	10	US-09-942-572-1	Sequence 1, Appli
c	325	16	0.5	1210	9	US-09-764-868-93	Sequence 93, Appl
c	326	16	0.5	1212	10	US-09-880-503-15	Sequence 15, Appl
c	327	16	0.5	1224	10	US-09-804-551B-23	Sequence 23, Appl
c	328	16	0.5	1230	10	US-09-770-445-35	Sequence 35, Appl
	329	16	0.5	1235	10	US-09-974-300-7279	Sequence 7279, Ap
c	330	16	0.5	1236	10	US-09-880-503-12	Sequence 12, Appl
	331	16	0.5	1254	10	US-09-965-553-25	Sequence 25, Appl
	332	16	0.5	1260	10	US-09-175-254-2	Sequence 2, Appli
	333	16	0.5	1260	10	US-09-870-379-1	Sequence 1, Appli
	334	16	0.5	1263	9	US-09-738-626-281	Sequence 281, App
c	335	16	0.5	1277	12	US-10-062-254-279	Sequence 279, App
	336	16	0.5	1283	10	US-09-925-302-307	Sequence 307, App
c	337	16	0.5	1287	10	US-09-741-669-201	Sequence 201, App
	338	16	0.5	1293	10	US-09-815-242-8128	Sequence 8128, Ap
	339	16	0.5	1293	10	US-09-815-242-8727	Sequence 8727, Ap
	340	16	0.5	1293	10	US-09-815-242-8952	Sequence 8952, Ap
c	341	16	0.5	1294	10	US-09-850-964-4	Sequence 4, Appli
c	342	16	0.5	1296	9	US-10-076-421-1	Sequence 1, Appli
	343	16	0.5	1305	10	US-09-815-242-4486	Sequence 4486, Ap
	344	16	0.5	1321	10	US-09-934-249-1	Sequence 1, Appli
	345	16	0.5	1393	9	US-09-841-730-24	Sequence 24, Appl
	346	16	0.5	1393	10	US-09-454-540-1	Sequence 1, Appli
	347	16	0.5	1393	10	US-09-859-894A-1	Sequence 1, Appli
	348	16	0.5	1403	9	US-10-098-841-280	Sequence 280, App
	349	16	0.5	1412	10	US-09-974-300-171	Sequence 171, App
	350	16	0.5	1422	9	US-09-738-626-3162	Sequence 3162, Ap
	351	16	0.5	1434	9	US-09-738-626-383	Sequence 383, App
	352	16	0.5	1434	9	US-10-047-412A-9	Sequence 9, Appli
	353	16	0.5	1434	9	US-10-047-412A-28	Sequence 28, Appl
c	354	16	0.5	1443	10	US-09-974-300-2435	Sequence 2435, Ap
	355	16	0.5	1457	9	US-09-954-531-982	Sequence 982, App
c	356	16	0.5	1475	10	US-09-735-705-122	Sequence 122, App
c	357	16	0.5	1475	10	US-09-850-716A-122	Sequence 122, App
c	358	16	0.5	1475	10	US-09-897-778-122	Sequence 122, App
	359	16	0.5	1476	9	US-09-974-298-94	Sequence 94, Appl
c	360	16	0.5	1511	9	US-10-078-650-11	Sequence 11, Appl
c	361	16	0.5	1519	10	US-09-392-822-2	Sequence 2, Appli
c	362	16	0.5	1523	9	US-09-938-842A-4708	Sequence 4708, Ap
c	363	16	0.5	1529	10	US-09-820-893-20	Sequence 20, Appl
c	364	16	0.5	1604	9	US-09-938-842A-3688	Sequence 3688, Ap
	365	16	0.5	1635	9	US-10-098-841-207	Sequence 207, App
c	366	16	0.5	1636	10	US-09-954-456-1185	Sequence 1185, Ap
c	367	16	0.5	1636	10	US-09-954-456-1850	Sequence 1850, Ap
	368	16	0.5	1637	10	US-09-925-301-353	Sequence 353, App
c	369	16	0.5	1674	9	US-09-935-726-1	Sequence 1, Appli
c	370	16	0.5	1674	12	US-10-127-551-1	Sequence 1, Appli
	371	16	0.5	1686	9	US-10-098-841-206	Sequence 206, App
	372	16	0.5	1693	10	US-09-867-550-1669	Sequence 1669, Ap
c	373	16	0.5	1695	10	US-09-764-878-428	Sequence 428, App
	374	16	0.5	1707	10	US-09-804-357-13	Sequence 13, Appl
	375	16	0.5	1707	10	US-09-804-006-13	Sequence 13, Appl
	376	16	0.5	1714	10	US-09-880-192-17	Sequence 17, Appl
	377	16	0.5	1720	10	US-09-746-359A-52	Sequence 52, Appl
c	378	16	0.5	1722	12	US-10-062-254-271	Sequence 271, App
	379	16	0.5	1732	10	US-09-923-556-1	Sequence 1, Appli
	380	16	0.5	1732	10	US-09-923-556-5	Sequence 5, Appli
	381	16	0.5	1740	9	US-09-922-364A-15	Sequence 15, Appl
	382	16	0.5	1740	9	US-09-254-590-15	Sequence 15, Appl
	383	16	0.5	1740	9	US-10-115-695-15	Sequence 15, Appl
	384	16	0.5	1749	9	US-09-938-842A-1851	Sequence 1851, Ap

	385	16	0.5	1750	9	US-10-028-072-187	Sequence 187, App
	386	16	0.5	1750	9	US-10-174-590-397	Sequence 397, App
	387	16	0.5	1750	9	US-10-176-758-397	Sequence 397, App
	388	16	0.5	1750	9	US-10-175-737-397	Sequence 397, App
	389	16	0.5	1750	12	US-10-052-586-397	Sequence 397, App
c	390	16	0.5	1779	10	US-09-833-381-1466	Sequence 1466, Ap
c	391	16	0.5	1779	10	US-09-833-381-1471	Sequence 1471, Ap
	392	16	0.5	1785	9	US-09-981-876-112	Sequence 112, App
	393	16	0.5	1791	10	US-09-880-107-3385	Sequence 3385, Ap
c	394	16	0.5	1792	9	US-09-954-531-1349	Sequence 1349, Ap
	395	16	0.5	1799	12	US-10-062-254-247	Sequence 247, App
	396	16	0.5	1801	10	US-09-746-359A-22	Sequence 22, Appl
c	397	16	0.5	1829	10	US-09-925-302-354	Sequence 354, App
c	398	16	0.5	1851	9	US-09-989-920-36	Sequence 36, Appl
	399	16	0.5	1890	10	US-09-880-107-3858	Sequence 3858, Ap
c	400	16	0.5	1891	10	US-09-969-708-146	Sequence 146, App
c	401	16	0.5	1893	10	US-09-815-242-8019	Sequence 8019, Ap
	402	16	0.5	1894	9	US-09-464-099A-66	Sequence 66, Appl
	403	16	0.5	1894	10	US-09-861-696-66	Sequence 66, Appl
c	404	16	0.5	1929	9	US-10-159-749-7	Sequence 7, Appli
c	405	16	0.5	1929	10	US-09-969-708-100	Sequence 100, App
c	406	16	0.5	1929	10	US-09-880-107-1636	Sequence 1636, Ap
	407	16	0.5	1931	10	US-09-925-301-454	Sequence 454, App
	408	16	0.5	1932	10	US-09-815-242-4615	Sequence 4615, Ap
	409	16	0.5	1935	10	US-09-815-242-8597	Sequence 8597, Ap
	410	16	0.5	1935	10	US-09-815-242-8900	Sequence 8900, Ap
	411	16	0.5	1951	9	US-09-736-968A-104	Sequence 104, App
c	412	16	0.5	1956	9	US-09-938-842A-2561	Sequence 2561, Ap
	413	16	0.5	1969	10	US-09-925-300-155	Sequence 155, App
c	414	16	0.5	1997	10	US-09-795-006A-21	Sequence 21, Appl
c	415	16	0.5	2000	9	US-09-938-842A-5058	Sequence 5058, Ap
c	416	16	0.5	2000	9	US-09-938-842A-5088	Sequence 5088, Ap
	417	16	0.5	2000	10	US-09-954-456-1237	Sequence 1237, Ap
	418	16	0.5	2000	10	US-09-997-165-3	Sequence 3, Appli
c	419	16	0.5	2000	10	US-09-887-576-837	Sequence 837, App
	420	16	0.5	2002	10	US-09-925-300-592	Sequence 592, App
c	421	16	0.5	2007	9	US-09-938-842A-2070	Sequence 2070, Ap
	422	16	0.5	2009	10	US-09-971-309-17	Sequence 17, Appl
	423	16	0.5	2025	9	US-09-738-626-2417	Sequence 2417, Ap
	424	16	0.5	2037	10	US-09-822-830A-25	Sequence 25, Appl
c	425	16	0.5	2056	10	US-09-925-301-482	Sequence 482, App
c	426	16	0.5	2066	10	US-09-835-996A-14	Sequence 14, Appl
	427	16	0.5	2130	9	US-09-740-369-1	Sequence 1, Appli
c	428	16	0.5	2136	9	US-09-981-876-120	Sequence 120, App
c	429	16	0.5	2136	10	US-09-820-893-44	Sequence 44, Appl
c	430	16	0.5	2143	10	US-09-833-381-1022	Sequence 1022, Ap
	431	16	0.5	2161	8	US-08-834-666A-17	Sequence 17, Appl
	432	16	0.5	2193	8	US-08-910-386A-15	Sequence 15, Appl
c	433	16	0.5	2196	10	US-09-925-300-504	Sequence 504, App
c	434	16	0.5	2230	10	US-09-925-299-87	Sequence 87, Appl
	435	16	0.5	2248	8	US-08-834-666A-15	Sequence 15, Appl
c	436	16	0.5	2261	10	US-09-940-919-3	Sequence 3, Appli
c	437	16	0.5	2261	10	US-09-917-800A-484	Sequence 484, App
	438	16	0.5	2279	10	US-09-954-456-269	Sequence 269, App
	439	16	0.5	2279	10	US-09-954-456-950	Sequence 950, App
	440	16	0.5	2279	10	US-09-954-456-1600	Sequence 1600, Ap
c	441	16	0.5	2289	10	US-09-815-242-7422	Sequence 7422, Ap
c	442	16	0.5	2289	10	US-09-850-964-3	Sequence 3, Appli
c	443	16	0.5	2294	10	US-09-735-705-123	Sequence 123, App
c	444	16	0.5	2294	10	US-09-850-716A-123	Sequence 123, App
c	445	16	0.5	2294	10	US-09-897-778-123	Sequence 123, App
c	446	16	0.5	2298	9	US-09-954-531-1386	Sequence 1386, Ap
c	447	16	0.5	2311	10	US-09-925-302-244	Sequence 244, App
c	448	16	0.5	2314	9	US-10-028-072-115	Sequence 115, App
c	449	16	0.5	2316	9	US-09-738-626-3161	Sequence 3161, Ap
c	450	16	0.5	2322	9	US-10-045-815-7	Sequence 7, Appli
	451	16	0.5	2335	10	US-09-742-684-3	Sequence 3, Appli
	452	16	0.5	2405	9	US-09-981-353-156	Sequence 156, App
c	453	16	0.5	2416	9	US-10-045-815-5	Sequence 5, Appli
	454	16	0.5	2418	9	US-09-938-842A-872	Sequence 872, App
	455	16	0.5	2435	10	US-09-823-038A-40	Sequence 40, Appl
	456	16	0.5	2453	12	US-10-005-858-1	Sequence 1, Appli
c	457	16	0.5	2466	12	US-10-044-090-251	Sequence 251, App
	458	16	0.5	2470	10	US-09-964-824A-555	Sequence 555, App
	459	16	0.5	2470	10	US-09-880-107-3310	Sequence 3310, Ap
c	460	16	0.5	2475	9	US-10-011-588-32	Sequence 32, Appl
	461	16	0.5	2475	9	US-10-028-072-195	Sequence 195, App
	462	16	0.5	2493	10	US-09-815-028-1	Sequence 1, Appli
c	463	16	0.5	2543	12	US-10-044-090-211	Sequence 211, App
c	464	16	0.5	2565	9	US-09-754-853A-5	Sequence 5, Appli
c	465	16	0.5	2565	9	US-09-754-853A-36	Sequence 36, Appl
c	466	16	0.5	2565	9	US-09-754-853A-37	Sequence 37, Appl
c	467	16	0.5	2565	9	US-09-754-853A-38	Sequence 38, Appl
c	468	16	0.5	2565	9	US-09-754-853A-39	Sequence 39, Appl
c	469	16	0.5	2565	9	US-09-754-853A-40	Sequence 40, Appl
c	470	16	0.5	2565	9	US-09-754-853A-41	Sequence 41, Appl
c	471	16	0.5	2565	9	US-09-754-853A-42	Sequence 42, Appl
c	472	16	0.5	2565	9	US-09-754-853A-43	Sequence 43, Appl
	473	16	0.5	2576	10	US-09-175-254-1	Sequence 1, Appli
c	474	16	0.5	2617	9	US-09-974-298-183	Sequence 183, App
	475	16	0.5	2622	10	US-09-950-370-3	Sequence 3, Appli
c	476	16	0.5	2634	9	US-09-754-853A-6	Sequence 6, Appli
c	477	16	0.5	2634	9	US-09-754-853A-28	Sequence 28, Appl
c	478	16	0.5	2634	9	US-09-754-853A-29	Sequence 29, Appl
c	479	16	0.5	2634	9	US-09-754-853A-30	Sequence 30, Appl
c	480	16	0.5	2634	9	US-09-754-853A-31	Sequence 31, Appl
c	481	16	0.5	2634	9	US-09-754-853A-32	Sequence 32, Appl
c	482	16	0.5	2634	9	US-09-754-853A-33	Sequence 33, Appl
c	483	16	0.5	2634	9	US-09-754-853A-34	Sequence 34, Appl
c	484	16	0.5	2634	9	US-09-754-853A-35	Sequence 35, Appl
c	485	16	0.5	2667	10	US-09-810-796-3	Sequence 3, Appli
c	486	16	0.5	2694	10	US-09-810-796-2	Sequence 2, Appli
c	487	16	0.5	2701	10	US-09-764-877-3378	Sequence 3378, Ap
	488	16	0.5	2706	10	US-09-950-370-13	Sequence 13, Appl
c	489	16	0.5	2708	12	US-10-044-090-122	Sequence 122, App
c	490	16	0.5	2720	9	US-10-174-590-137	Sequence 137, App
c	491	16	0.5	2720	9	US-10-176-758-137	Sequence 137, App
c	492	16	0.5	2720	9	US-10-175-737-137	Sequence 137, App
c	493	16	0.5	2720	10	US-09-079-892-5	Sequence 5, Appli
c	494	16	0.5	2720	12	US-10-052-586-137	Sequence 137, App
	495	16	0.5	2725	10	US-09-962-436-40	Sequence 40, Appl
	496	16	0.5	2730	9	US-09-935-371-39	Sequence 39, Appl
c	497	16	0.5	2733	10	US-09-789-561-59	Sequence 59, Appl
c	498	16	0.5	2739	10	US-09-864-864-287	Sequence 287, App
c	499	16	0.5	2739	10	US-09-864-864-322	Sequence 322, App
c	500	16	0.5	2772	10	US-09-825-147-1	Sequence 1, Appli
	501	16	0.5	2784	9	US-10-071-766-42	Sequence 42, Appl
	502	16	0.5	2814	9	US-09-935-371-38	Sequence 38, Appl
c	503	16	0.5	2832	12	US-10-115-178-2	Sequence 2, Appli
c	504	16	0.5	2838	10	US-09-954-456-700	Sequence 700, App
c	505	16	0.5	2838	10	US-09-954-456-966	Sequence 966, App
c	506	16	0.5	2838	10	US-09-954-456-1143	Sequence 1143, Ap
	507	16	0.5	2843	9	US-09-854-133-430	Sequence 430, App
	508	16	0.5	2843	10	US-09-738-973-430	Sequence 430, App
c	509	16	0.5	2853	10	US-09-905-983-45	Sequence 45, Appl
	510	16	0.5	2883	9	US-10-174-590-157	Sequence 157, App
	511	16	0.5	2883	9	US-10-176-758-157	Sequence 157, App
	512	16	0.5	2883	9	US-10-175-737-157	Sequence 157, App
	513	16	0.5	2883	12	US-10-052-586-157	Sequence 157, App
	514	16	0.5	2910	10	US-09-833-790-418	Sequence 418, App
	515	16	0.5	2910	10	US-09-880-107-3784	Sequence 3784, Ap
	516	16	0.5	2934	10	US-09-815-242-4326	Sequence 4326, Ap
c	517	16	0.5	2935	9	US-10-047-412A-27	Sequence 27, Appl
	518	16	0.5	2936	9	US-10-108-605-54	Sequence 54, Appl
	519	16	0.5	2956	9	US-09-978-295A-89	Sequence 89, Appl
	520	16	0.5	2956	9	US-09-978-697-89	Sequence 89, Appl
	521	16	0.5	2956	9	US-09-978-192A-89	Sequence 89, Appl
	522	16	0.5	2956	9	US-09-999-832A-89	Sequence 89, Appl
	523	16	0.5	2956	9	US-09-978-189-89	Sequence 89, Appl
	524	16	0.5	2956	9	US-10-174-590-73	Sequence 73, Appl
	525	16	0.5	2956	9	US-10-176-758-73	Sequence 73, Appl
	526	16	0.5	2956	9	US-10-175-737-73	Sequence 73, Appl
	527	16	0.5	2956	12	US-10-052-586-73	Sequence 73, Appl
	528	16	0.5	2964	10	US-09-815-242-8094	Sequence 8094, Ap
	529	16	0.5	2981	10	US-09-950-370-1	Sequence 1, Appli
c	530	16	0.5	3016	10	US-09-925-301-411	Sequence 411, App

	531	16	0.5	3024	10	US-09-833-381-889	Sequence 889, App
	532	16	0.5	3035	10	US-09-864-864-311	Sequence 311, App
c	533	16	0.5	3059	10	US-09-770-595A-2	Sequence 2, Appli
	534	16	0.5	3065	10	US-09-950-370-11	Sequence 11, Appl
c	535	16	0.5	3071	10	US-09-810-796-1	Sequence 1, Appli
c	536	16	0.5	3074	10	US-09-813-148-1	Sequence 1, Appli
c	537	16	0.5	3111	10	US-09-825-147-3	Sequence 3, Appli
c	538	16	0.5	3124	9	US-09-974-298-128	Sequence 128, App
	539	16	0.5	3191	9	US-09-989-920-54	Sequence 54, Appl
c	540	16	0.5	3222	10	US-09-974-300-2654	Sequence 2654, Ap
c	541	16	0.5	3274	10	US-09-070-927A-350	Sequence 350, App
	542	16	0.5	3313	9	US-09-978-295A-389	Sequence 389, App
	543	16	0.5	3313	9	US-09-978-697-389	Sequence 389, App
	544	16	0.5	3313	9	US-09-978-192A-389	Sequence 389, App
	545	16	0.5	3313	9	US-09-999-832A-389	Sequence 389, App
	546	16	0.5	3313	9	US-09-978-189-389	Sequence 389, App
	547	16	0.5	3313	9	US-10-174-590-77	Sequence 77, Appl
	548	16	0.5	3313	9	US-10-176-758-77	Sequence 77, Appl
	549	16	0.5	3313	9	US-10-175-737-77	Sequence 77, Appl
	550	16	0.5	3313	12	US-10-052-586-77	Sequence 77, Appl
c	551	16	0.5	3358	9	US-09-954-531-1354	Sequence 1354, Ap
	552	16	0.5	3376	10	US-09-834-975-775	Sequence 775, App
	553	16	0.5	3400	10	US-09-987-025-1	Sequence 1, Appli
c	554	16	0.5	3447	10	US-09-815-242-9728	Sequence 9728, Ap
	555	16	0.5	3516	10	US-09-746-359A-10	Sequence 10, Appl
	556	16	0.5	3609	10	US-09-799-875-6	Sequence 6, Appli
	557	16	0.5	3637	10	US-09-910-150-1	Sequence 1, Appli
c	558	16	0.5	3641	9	US-10-078-650-1	Sequence 1, Appli
c	559	16	0.5	3657	10	US-09-728-422-4	Sequence 4, Appli
	560	16	0.5	3687	10	US-09-790-318-3	Sequence 3, Appli
	561	16	0.5	3803	9	US-10-114-893-179	Sequence 179, App
	562	16	0.5	3830	9	US-10-078-770-185	Sequence 185, App
	563	16	0.5	3920	10	US-09-923-684-1	Sequence 1, Appli
	564	16	0.5	3954	10	US-09-815-242-4063	Sequence 4063, Ap
c	565	16	0.5	3991	12	US-10-044-090-666	Sequence 666, App
	566	16	0.5	4098	10	US-09-962-436-37	Sequence 37, Appl
	567	16	0.5	4134	9	US-09-975-553-17	Sequence 17, Appl
	568	16	0.5	4286	10	US-09-899-634A-1	Sequence 1, Appli
	569	16	0.5	4339	12	US-10-002-600-50	Sequence 50, Appl
	570	16	0.5	4368	9	US-09-981-353-15	Sequence 15, Appl
c	571	16	0.5	4501	9	US-09-754-853A-10	Sequence 10, Appl
c	572	16	0.5	4501	9	US-09-754-853A-11	Sequence 11, Appl
c	573	16	0.5	4508	9	US-09-754-853A-18	Sequence 18, Appl
c	574	16	0.5	4508	9	US-09-754-853A-19	Sequence 19, Appl
c	575	16	0.5	4508	9	US-09-754-853A-20	Sequence 20, Appl
c	576	16	0.5	4508	9	US-09-754-853A-21	Sequence 21, Appl
c	577	16	0.5	4508	9	US-09-754-853A-22	Sequence 22, Appl
c	578	16	0.5	4508	9	US-09-754-853A-23	Sequence 23, Appl
c	579	16	0.5	4522	9	US-09-754-853A-12	Sequence 12, Appl
c	580	16	0.5	4522	9	US-09-754-853A-13	Sequence 13, Appl
c	581	16	0.5	4523	9	US-09-754-853A-8	Sequence 8, Appli
c	582	16	0.5	4523	9	US-09-754-853A-9	Sequence 9, Appli
c	583	16	0.5	4523	9	US-09-754-853A-14	Sequence 14, Appl
c	584	16	0.5	4523	9	US-09-754-853A-15	Sequence 15, Appl
c	585	16	0.5	4523	9	US-09-754-853A-16	Sequence 16, Appl
c	586	16	0.5	4523	9	US-09-754-853A-17	Sequence 17, Appl
	587	16	0.5	4600	9	US-09-736-457-1797	Sequence 1797, Ap
	588	16	0.5	4600	9	US-09-902-941-1797	Sequence 1797, Ap
	589	16	0.5	4600	9	US-09-849-626-1797	Sequence 1797, Ap
	590	16	0.5	4739	9	US-09-954-531-586	Sequence 586, App
c	591	16	0.5	4780	10	US-09-962-436-287	Sequence 287, App
	592	16	0.5	5025	10	US-09-960-253-176	Sequence 176, App
	593	16	0.5	5222	10	US-09-416-384A-1	Sequence 1, Appli
	594	16	0.5	5427	9	US-09-981-353-159	Sequence 159, App
	595	16	0.5	5566	10	US-09-416-384A-4	Sequence 4, Appli
	596	16	0.5	5597	10	US-09-895-072-4	Sequence 4, Appli
	597	16	0.5	5597	10	US-09-986-552-4	Sequence 4, Appli
	598	16	0.5	5756	9	US-09-548-880-1	Sequence 1, Appli
	599	16	0.5	5983	10	US-09-799-875-4	Sequence 4, Appli
c	600	16	0.5	6057	12	US-10-044-090-239	Sequence 239, App
c	601	16	0.5	6181	9	US-09-764-868-1280	Sequence 1280, Ap
c	602	16	0.5	6356	10	US-09-770-595A-1	Sequence 1, Appli
	603	16	0.5	6693	9	US-10-071-766-135	Sequence 135, App
c	604	16	0.5	6751	10	US-09-761-466-5	Sequence 5, Appli
c	605	16	0.5	8119	10	US-09-802-669-45	Sequence 45, Appl
c	606	16	0.5	8372	9	US-09-902-941-1893	Sequence 1893, Ap
c	607	16	0.5	8372	9	US-09-849-626-1893	Sequence 1893, Ap
	608	16	0.5	9567	12	US-10-016-768-12	Sequence 12, Appl
	609	16	0.5	9574	12	US-10-016-768-9	Sequence 9, Appli
	610	16	0.5	9581	10	US-09-764-860-895	Sequence 895, App
	611	16	0.5	10883	10	US-09-853-386-13	Sequence 13, Appl
	612	16	0.5	11186	10	US-09-957-997-1	Sequence 1, Appli
	613	16	0.5	11580	10	US-09-880-107-3436	Sequence 3436, Ap
c	614	16	0.5	11945	9	US-09-922-549B-67	Sequence 67, Appl
c	615	16	0.5	12606	10	US-09-957-974-2	Sequence 2, Appli
c	616	16	0.5	12728	9	US-09-922-549B-65	Sequence 65, Appl
	617	16	0.5	13341	8	US-08-910-386A-1	Sequence 1, Appli
c	618	16	0.5	14175	9	US-09-764-868-1474	Sequence 1474, Ap
c	619	16	0.5	15613	10	US-09-764-877-3525	Sequence 3525, Ap
c	620	16	0.5	16545	10	US-09-835-996A-32	Sequence 32, Appl
	621	16	0.5	17379	10	US-09-764-860-861	Sequence 861, App
	622	16	0.5	31412	9	US-10-109-551-3	Sequence 3, Appli
	623	16	0.5	41936	10	US-09-967-768A-116	Sequence 116, App
	624	16	0.5	42999	9	US-09-799-462A-17	Sequence 17, Appl
c	625	16	0.5	42999	9	US-09-799-462A-17	Sequence 17, Appl
	626	16	0.5	42999	9	US-10-125-767-17	Sequence 17, Appl
c	627	16	0.5	42999	9	US-10-125-767-17	Sequence 17, Appl
	628	16	0.5	78056	9	US-10-109-551-1	Sequence 1, Appli
	629	16	0.5	80959	9	US-09-858-546-3	Sequence 3, Appli
c	630	16	0.5	167343	10	US-09-962-436-281	Sequence 281, App
c	631	16	0.5	167343	10	US-09-964-824A-273	Sequence 273, App
	632	16	0.5	174424	10	US-09-967-768A-314	Sequence 314, App
	633	16	0.5	249487	9	US-10-026-188-3	Sequence 3, Appli
	634	16	0.5	335913	9	US-09-754-853A-2	Sequence 2, Appli
	635	16	0.5	335913	9	US-09-754-853A-3	Sequence 3, Appli
c	636	16	0.5	368004	10	US-09-949-654-3	Sequence 3, Appli
c	637	16	0.5	659158	9	US-09-771-208-20	Sequence 20, Appl
	638	16	0.5	684973	10	US-09-263-959-1	Sequence 1, Appli
	639	16	0.5	1503841	9	US-09-946-807-1	Sequence 1, Appli
c	640	16	0.5	1503841	9	US-09-946-807-1	Sequence 1, Appli
	641	16	0.5	1503841	10	US-09-795-668-1	Sequence 1, Appli
c	642	16	0.5	1503841	10	US-09-795-668-1	Sequence 1, Appli
	643	16	0.5	1503841	10	US-09-795-686-1	Sequence 1, Appli
c	644	16	0.5	1503841	10	US-09-795-686-1	Sequence 1, Appli
	645	16	0.5	3309400	9	US-09-738-626-1	Sequence 1, Appli
c	646	15	0.5	21	10	US-09-828-034-10	Sequence 10, Appl
c	647	15	0.5	30	9	US-10-011-635A-4	Sequence 4, Appli
	648	15	0.5	50	10	US-09-782-650-16	Sequence 16, Appl
c	649	15	0.5	51	10	US-09-782-650-17	Sequence 17, Appl
	650	15	0.5	80	10	US-09-864-761-27458	Sequence 27458, A
	651	15	0.5	92	10	US-09-864-761-21237	Sequence 21237, A
c	652	15	0.5	113	9	US-09-764-868-1360	Sequence 1360, Ap
c	653	15	0.5	118	10	US-09-770-696-700	Sequence 700, App
c	654	15	0.5	119	10	US-09-865-499-4	Sequence 4, Appli
c	655	15	0.5	122	10	US-09-864-761-25491	Sequence 25491, A
	656	15	0.5	129	10	US-09-923-876-1995	Sequence 1995, Ap
c	657	15	0.5	131	10	US-09-998-598-982	Sequence 982, App
	658	15	0.5	133	10	US-09-923-876-2863	Sequence 2863, Ap
c	659	15	0.5	134	10	US-09-983-965-2662	Sequence 2662, Ap
	660	15	0.5	138	10	US-09-244-694-129	Sequence 129, App
	661	15	0.5	143	10	US-09-864-761-26862	Sequence 26862, A
	662	15	0.5	143	10	US-09-974-300-5308	Sequence 5308, Ap
c	663	15	0.5	147	10	US-09-924-035A-247	Sequence 247, App
	664	15	0.5	149	10	US-09-878-574-14501	Sequence 14501, A
c	665	15	0.5	155	10	US-09-770-696-446	Sequence 446, App
	666	15	0.5	159	9	US-09-738-626-3296	Sequence 3296, Ap
c	667	15	0.5	161	10	US-09-884-441-35	Sequence 35, Appl
c	668	15	0.5	168	10	US-09-879-536-222	Sequence 222, App
	669	15	0.5	172	10	US-09-878-574-12516	Sequence 12516, A
	670	15	0.5	175	10	US-09-954-456-1477	Sequence 1477, Ap
	671	15	0.5	184	10	US-09-864-761-31978	Sequence 31978, A
c	672	15	0.5	194	10	US-09-864-761-32733	Sequence 32733, A
c	673	15	0.5	199	10	US-09-864-761-22166	Sequence 22166, A
	674	15	0.5	203	10	US-09-294-093B-4713	Sequence 4713, Ap
c	675	15	0.5	203	10	US-09-878-574-8445	Sequence 8445, Ap
c	676	15	0.5	204	12	US-10-010-901-10	Sequence 10, Appl

	677	15	0.5	207	10	US-09-974-300-4196	Sequence 4196, Ap
c	678	15	0.5	213	10	US-09-923-876-178	Sequence 178, App
	679	15	0.5	215	10	US-09-815-242-1079	Sequence 1079, Ap
c	680	15	0.5	216	9	US-09-796-692-7248	Sequence 7248, Ap
c	681	15	0.5	217	10	US-09-878-574-12062	Sequence 12062, A
c	682	15	0.5	217	10	US-09-962-832-74	Sequence 74, Appl
c	683	15	0.5	222	10	US-09-974-300-4139	Sequence 4139, Ap
	684	15	0.5	226	10	US-09-728-445-159	Sequence 159, App
c	685	15	0.5	228	10	US-09-815-242-819	Sequence 819, App
	686	15	0.5	231	10	US-09-878-574-8627	Sequence 8627, Ap
c	687	15	0.5	231	10	US-09-880-107-1352	Sequence 1352, Ap
	688	15	0.5	233	10	US-09-960-352-6124	Sequence 6124, Ap
	689	15	0.5	238	10	US-09-983-965-3902	Sequence 3902, Ap
c	690	15	0.5	246	10	US-09-878-574-6208	Sequence 6208, Ap
	691	15	0.5	247	10	US-09-923-876-2971	Sequence 2971, Ap
	692	15	0.5	249	10	US-09-783-590-10662	Sequence 10662, A
c	693	15	0.5	249	10	US-09-883-825-40	Sequence 40, Appl
c	694	15	0.5	251	9	US-09-933-797-319	Sequence 319, App
c	695	15	0.5	251	10	US-09-878-574-7980	Sequence 7980, Ap
	696	15	0.5	252	10	US-09-923-876-496	Sequence 496, App
c	697	15	0.5	252	10	US-09-878-574-15770	Sequence 15770, A
	698	15	0.5	254	10	US-09-878-574-10714	Sequence 10714, A
c	699	15	0.5	254	10	US-09-878-574-11391	Sequence 11391, A
	700	15	0.5	255	9	US-09-738-626-666	Sequence 666, App
	701	15	0.5	258	10	US-09-923-876-1914	Sequence 1914, Ap
c	702	15	0.5	258	10	US-09-878-574-10667	Sequence 10667, A
c	703	15	0.5	259	10	US-09-923-876-1666	Sequence 1666, Ap
c	704	15	0.5	259	10	US-09-912-020-471	Sequence 471, App
c	705	15	0.5	260	10	US-09-912-020-67	Sequence 67, Appl
	706	15	0.5	260	10	US-09-878-574-11123	Sequence 11123, A
c	707	15	0.5	260	10	US-09-878-574-13054	Sequence 13054, A
	708	15	0.5	261	10	US-09-923-876-695	Sequence 695, App
c	709	15	0.5	261	10	US-09-864-761-17969	Sequence 17969, A
	710	15	0.5	262	10	US-09-923-876-3431	Sequence 3431, Ap
c	711	15	0.5	264	10	US-09-294-093B-730	Sequence 730, App
c	712	15	0.5	265	9	US-10-040-739-1219	Sequence 1219, Ap
c	713	15	0.5	265	10	US-09-294-093B-792	Sequence 792, App
	714	15	0.5	265	10	US-09-878-574-9152	Sequence 9152, Ap
	715	15	0.5	267	10	US-09-923-876-5854	Sequence 5854, Ap
c	716	15	0.5	268	10	US-09-880-107-2915	Sequence 2915, Ap
	717	15	0.5	271	10	US-09-923-876-5076	Sequence 5076, Ap
	718	15	0.5	271	10	US-09-878-574-7633	Sequence 7633, Ap
	719	15	0.5	273	10	US-09-878-574-8991	Sequence 8991, Ap
	720	15	0.5	273	10	US-09-878-574-15635	Sequence 15635, A
c	721	15	0.5	274	10	US-09-878-574-9369	Sequence 9369, Ap
	722	15	0.5	276	10	US-09-793-306-151	Sequence 151, App
c	723	15	0.5	276	10	US-09-960-352-11026	Sequence 11026, A
	724	15	0.5	280	10	US-09-923-876-6034	Sequence 6034, Ap
c	725	15	0.5	280	10	US-09-878-574-8050	Sequence 8050, Ap
	726	15	0.5	281	10	US-09-294-093B-5703	Sequence 5703, Ap
c	727	15	0.5	281	10	US-09-878-574-15012	Sequence 15012, A
	728	15	0.5	283	10	US-09-294-093B-931	Sequence 931, App
	729	15	0.5	283	10	US-09-878-574-5469	Sequence 5469, Ap
c	730	15	0.5	284	10	US-09-867-701-2907	Sequence 2907, Ap
	731	15	0.5	285	10	US-09-912-020-191	Sequence 191, App
	732	15	0.5	288	10	US-09-783-590-4734	Sequence 4734, Ap
c	733	15	0.5	288	10	US-09-867-701-5557	Sequence 5557, Ap
	734	15	0.5	289	10	US-09-294-093B-5584	Sequence 5584, Ap
c	735	15	0.5	290	10	US-09-294-093B-243	Sequence 243, App
	736	15	0.5	294	10	US-09-244-694-137	Sequence 137, App
	737	15	0.5	294	10	US-09-864-761-19936	Sequence 19936, A
c	738	15	0.5	294	10	US-09-783-590-4	Sequence 4, Appli
	739	15	0.5	296	10	US-09-294-093B-6090	Sequence 6090, Ap
c	740	15	0.5	297	10	US-09-294-093B-1856	Sequence 1856, Ap
	741	15	0.5	300	10	US-09-294-093B-3923	Sequence 3923, Ap
	742	15	0.5	300	10	US-09-941-997-1	Sequence 1, Appli
c	743	15	0.5	301	9	US-09-796-692-6537	Sequence 6537, Ap
	744	15	0.5	302	10	US-09-867-701-7885	Sequence 7885, Ap
	745	15	0.5	304	9	US-09-928-457-69	Sequence 69, Appl
	746	15	0.5	305	9	US-10-040-739-844	Sequence 844, App
c	747	15	0.5	306	10	US-09-815-242-1364	Sequence 1364, Ap
c	748	15	0.5	306	10	US-09-815-242-1523	Sequence 1523, Ap
c	749	15	0.5	306	10	US-09-815-242-1649	Sequence 1649, Ap
	750	15	0.5	307	10	US-09-867-701-10620	Sequence 10620, A
	751	15	0.5	311	10	US-09-960-352-14967	Sequence 14967, A
c	752	15	0.5	322	10	US-09-960-352-6429	Sequence 6429, Ap
	753	15	0.5	327	10	US-09-880-107-783	Sequence 783, App
c	754	15	0.5	330	10	US-09-815-242-2247	Sequence 2247, Ap
	755	15	0.5	332	10	US-09-294-093B-2203	Sequence 2203, Ap
	756	15	0.5	337	9	US-10-015-219-330	Sequence 330, App
	757	15	0.5	337	10	US-09-777-564-330	Sequence 330, App
	758	15	0.5	338	9	US-09-933-797-130	Sequence 130, App
c	759	15	0.5	339	9	US-09-920-455-161	Sequence 161, App
c	760	15	0.5	342	10	US-09-960-352-1348	Sequence 1348, Ap
c	761	15	0.5	342	10	US-09-960-352-13723	Sequence 13723, A
c	762	15	0.5	343	10	US-09-960-352-14433	Sequence 14433, A
	763	15	0.5	345	10	US-09-878-574-863	Sequence 863, App
	764	15	0.5	346	9	US-09-925-664-50	Sequence 50, Appl
c	765	15	0.5	348	10	US-09-960-352-13719	Sequence 13719, A
	766	15	0.5	349	9	US-09-996-140-94	Sequence 94, Appl
	767	15	0.5	350	10	US-09-983-965-840	Sequence 840, App
c	768	15	0.5	351	10	US-09-867-701-4299	Sequence 4299, Ap
	769	15	0.5	353	10	US-09-960-352-12027	Sequence 12027, A
	770	15	0.5	355	10	US-09-867-701-4602	Sequence 4602, Ap
	771	15	0.5	362	10	US-09-244-694-138	Sequence 138, App
c	772	15	0.5	363	10	US-09-878-574-391	Sequence 391, App
c	773	15	0.5	366	10	US-09-960-352-3994	Sequence 3994, Ap
c	774	15	0.5	367	10	US-09-867-701-8008	Sequence 8008, Ap
	775	15	0.5	369	10	US-09-960-352-10070	Sequence 10070, A
	776	15	0.5	369	10	US-09-983-965-3566	Sequence 3566, Ap
	777	15	0.5	371	10	US-09-960-352-8510	Sequence 8510, Ap
c	778	15	0.5	373	10	US-09-878-574-2299	Sequence 2299, Ap
	779	15	0.5	375	10	US-09-867-701-3254	Sequence 3254, Ap
	780	15	0.5	376	9	US-09-796-692-3399	Sequence 3399, Ap
	781	15	0.5	378	10	US-09-878-574-1986	Sequence 1986, Ap
	782	15	0.5	381	10	US-09-770-791-173	Sequence 173, App
	783	15	0.5	381	10	US-09-983-965-5851	Sequence 5851, Ap
c	784	15	0.5	381	10	US-09-999-256-12	Sequence 12, Appl
	785	15	0.5	382	10	US-09-880-107-2794	Sequence 2794, Ap
c	786	15	0.5	383	10	US-09-867-701-5594	Sequence 5594, Ap
c	787	15	0.5	384	10	US-09-960-352-8976	Sequence 8976, Ap
c	788	15	0.5	386	10	US-09-960-352-11656	Sequence 11656, A
c	789	15	0.5	388	10	US-09-347-331-5	Sequence 5, Appli
c	790	15	0.5	388	10	US-09-878-574-528	Sequence 528, App
	791	15	0.5	388	10	US-09-867-701-4534	Sequence 4534, Ap
	792	15	0.5	391	10	US-09-770-791-9	Sequence 9, Appli
	793	15	0.5	392	10	US-09-880-107-353	Sequence 353, App
	794	15	0.5	393	10	US-09-920-300A-1581	Sequence 1581, Ap
	795	15	0.5	393	12	US-10-033-528-1581	Sequence 1581, Ap
c	796	15	0.5	396	10	US-09-864-761-5393	Sequence 5393, Ap
	797	15	0.5	396	10	US-09-960-352-870	Sequence 870, App
	798	15	0.5	397	10	US-09-974-300-3551	Sequence 3551, Ap
	799	15	0.5	398	10	US-09-960-352-12334	Sequence 12334, A
	800	15	0.5	398	10	US-09-983-965-184	Sequence 184, App
	801	15	0.5	401	10	US-09-920-345-24	Sequence 24, Appl
	802	15	0.5	402	10	US-09-974-300-3122	Sequence 3122, Ap
c	803	15	0.5	404	10	US-09-878-574-3908	Sequence 3908, Ap
c	804	15	0.5	405	10	US-09-983-965-5418	Sequence 5418, Ap
	805	15	0.5	406	10	US-09-783-590-11317	Sequence 11317, A
	806	15	0.5	406	10	US-09-960-352-1716	Sequence 1716, Ap
c	807	15	0.5	406	10	US-09-949-192-20	Sequence 20, Appl
c	808	15	0.5	407	9	US-09-933-797-759	Sequence 759, App
	809	15	0.5	407	9	US-10-025-380-174	Sequence 174, App
	810	15	0.5	407	10	US-09-922-217-174	Sequence 174, App
	811	15	0.5	407	10	US-09-833-263-174	Sequence 174, App
c	812	15	0.5	408	10	US-09-878-574-1046	Sequence 1046, Ap
	813	15	0.5	410	10	US-09-783-590-7404	Sequence 7404, Ap
	814	15	0.5	410	10	US-09-960-352-8643	Sequence 8643, Ap
c	815	15	0.5	411	9	US-10-040-739-599	Sequence 599, App
	816	15	0.5	411	9	US-09-854-133-361	Sequence 361, App
	817	15	0.5	411	10	US-09-738-973-361	Sequence 361, App
c	818	15	0.5	412	10	US-09-960-352-8231	Sequence 8231, Ap
c	819	15	0.5	413	10	US-09-919-580-301	Sequence 301, App
c	820	15	0.5	413	10	US-09-960-352-138	Sequence 138, App
c	821	15	0.5	413	10	US-09-960-352-1603	Sequence 1603, Ap
	822	15	0.5	414	10	US-09-815-242-1060	Sequence 1060, Ap

c 823	15	0.5	414	10	US-09-960-352-13586	Sequence 13586, A
c 824	15	0.5	415	10	US-09-783-590-7545	Sequence 7545, Ap
825	15	0.5	415	10	US-09-960-352-7067	Sequence 7067, Ap
826	15	0.5	417	10	US-09-867-550-731	Sequence 731, App
c 827	15	0.5	419	10	US-09-878-574-4107	Sequence 4107, Ap
c 828	15	0.5	421	10	US-09-960-352-3158	Sequence 3158, Ap
829	15	0.5	421	10	US-09-983-965-4281	Sequence 4281, Ap
830	15	0.5	422	10	US-09-864-761-10577	Sequence 10577, A
c 831	15	0.5	423	10	US-09-974-300-7159	Sequence 7159, Ap
832	15	0.5	424	9	US-09-938-842A-3801	Sequence 3801, Ap
c 833	15	0.5	424	10	US-09-960-352-14599	Sequence 14599, A
c 834	15	0.5	425	9	US-10-114-893-91	Sequence 91, Appl
c 835	15	0.5	425	10	US-09-876-889-249	Sequence 249, App
836	15	0.5	426	10	US-09-764-877-152	Sequence 152, App
c 837	15	0.5	427	10	US-09-960-352-6081	Sequence 6081, Ap
838	15	0.5	429	10	US-09-878-574-3832	Sequence 3832, Ap
839	15	0.5	430	10	US-09-960-352-12560	Sequence 12560, A
c 840	15	0.5	430	10	US-09-983-965-2989	Sequence 2989, Ap
841	15	0.5	433	9	US-10-136-224-11	Sequence 11, Appl
842	15	0.5	433	10	US-09-960-352-7845	Sequence 7845, Ap
843	15	0.5	433	10	US-09-983-965-3790	Sequence 3790, Ap
c 844	15	0.5	435	10	US-09-867-701-7104	Sequence 7104, Ap
c 845	15	0.5	435	10	US-09-960-352-14799	Sequence 14799, A
846	15	0.5	435	10	US-09-880-107-213	Sequence 213, App
c 847	15	0.5	436	10	US-09-728-445-698	Sequence 698, App
848	15	0.5	436	10	US-09-983-965-498	Sequence 498, App
849	15	0.5	438	10	US-09-960-352-9523	Sequence 9523, Ap
c 850	15	0.5	438	10	US-09-983-965-2960	Sequence 2960, Ap
c 851	15	0.5	441	10	US-09-960-352-11772	Sequence 11772, A
c 852	15	0.5	442	9	US-09-954-531-914	Sequence 914, App
c 853	15	0.5	442	9	US-09-954-531-1324	Sequence 1324, Ap
c 854	15	0.5	442	9	US-09-764-868-85	Sequence 85, Appl
c 855	15	0.5	444	10	US-09-770-444-835	Sequence 835, App
856	15	0.5	444	10	US-09-864-761-3157	Sequence 3157, Ap
857	15	0.5	444	10	US-09-974-300-3588	Sequence 3588, Ap
c 858	15	0.5	445	10	US-09-954-456-221	Sequence 221, App
c 859	15	0.5	445	10	US-09-954-456-1362	Sequence 1362, Ap
860	15	0.5	446	9	US-10-015-219-423	Sequence 423, App
861	15	0.5	446	10	US-09-777-564-423	Sequence 423, App
862	15	0.5	447	9	US-10-015-219-139	Sequence 139, App
863	15	0.5	447	10	US-09-777-564-139	Sequence 139, App
864	15	0.5	447	10	US-09-864-761-2055	Sequence 2055, Ap
865	15	0.5	450	9	US-09-854-133-145	Sequence 145, App
c 866	15	0.5	450	10	US-09-770-444-704	Sequence 704, App
867	15	0.5	450	10	US-09-738-973-145	Sequence 145, App
c 868	15	0.5	450	10	US-09-960-352-5068	Sequence 5068, Ap
869	15	0.5	451	9	US-09-796-692-3075	Sequence 3075, Ap
c 870	15	0.5	451	10	US-09-867-550-1673	Sequence 1673, Ap
871	15	0.5	452	9	US-09-796-692-4129	Sequence 4129, Ap
c 872	15	0.5	452	10	US-09-770-444-645	Sequence 645, App
873	15	0.5	453	10	US-09-864-761-4489	Sequence 4489, Ap
874	15	0.5	453	10	US-09-864-761-10816	Sequence 10816, A
875	15	0.5	453	10	US-09-867-701-1871	Sequence 1871, Ap
c 876	15	0.5	455	10	US-09-960-352-7568	Sequence 7568, Ap
c 877	15	0.5	456	10	US-09-783-590-3028	Sequence 3028, Ap
c 878	15	0.5	457	9	US-09-764-868-504	Sequence 504, App
879	15	0.5	458	9	US-09-920-455-171	Sequence 171, App
c 880	15	0.5	459	12	US-10-027-559-1	Sequence 1, Appli
881	15	0.5	460	10	US-09-833-790-185	Sequence 185, App
882	15	0.5	463	9	US-09-954-531-868	Sequence 868, App
883	15	0.5	463	10	US-09-954-456-911	Sequence 911, App
c 884	15	0.5	464	10	US-09-960-352-1254	Sequence 1254, Ap
885	15	0.5	464	10	US-09-925-300-659	Sequence 659, App
886	15	0.5	465	9	US-09-938-842A-276	Sequence 276, App
c 887	15	0.5	468	10	US-09-864-761-3022	Sequence 3022, Ap
888	15	0.5	469	10	US-09-770-444-219	Sequence 219, App
889	15	0.5	469	10	US-09-867-701-2855	Sequence 2855, Ap
c 890	15	0.5	469	10	US-09-867-701-6813	Sequence 6813, Ap
891	15	0.5	470	10	US-09-864-761-3999	Sequence 3999, Ap
c 892	15	0.5	472	10	US-09-917-800A-662	Sequence 662, App
893	15	0.5	473	10	US-09-790-099-7	Sequence 7, Appli
894	15	0.5	473	10	US-09-864-761-15464	Sequence 15464, A
c 895	15	0.5	475	10	US-09-934-249-15	Sequence 15, Appl
c 896	15	0.5	479	10	US-09-864-761-1205	Sequence 1205, Ap
897	15	0.5	479	10	US-09-864-761-10224	Sequence 10224, A
c 898	15	0.5	480	10	US-09-864-761-133	Sequence 133, App
c 899	15	0.5	481	10	US-09-864-761-15122	Sequence 15122, A
c 900	15	0.5	484	9	US-09-796-692-9286	Sequence 9286, Ap
c 901	15	0.5	484	10	US-09-560-863-982	Sequence 982, App
c 902	15	0.5	486	10	US-09-864-761-14504	Sequence 14504, A
c 903	15	0.5	486	10	US-09-682-706-4	Sequence 4, Appli
904	15	0.5	486	10	US-09-917-800A-1171	Sequence 1171, Ap
c 905	15	0.5	489	10	US-09-244-694-153	Sequence 153, App
c 906	15	0.5	498	10	US-09-783-590-7412	Sequence 7412, Ap
907	15	0.5	498	10	US-09-778-844-188	Sequence 188, App
908	15	0.5	500	10	US-09-917-800A-1273	Sequence 1273, Ap
c 909	15	0.5	501	10	US-09-864-761-8824	Sequence 8824, Ap
910	15	0.5	501	10	US-09-864-761-16425	Sequence 16425, A
c 911	15	0.5	501	10	US-09-884-441-118	Sequence 118, App
912	15	0.5	501	10	US-09-833-381-133	Sequence 133, App
c 913	15	0.5	502	10	US-09-783-590-5959	Sequence 5959, Ap
c 914	15	0.5	505	10	US-09-783-590-12385	Sequence 12385, A
915	15	0.5	506	10	US-09-895-828-75	Sequence 75, Appl
c 916	15	0.5	508	10	US-09-783-590-3069	Sequence 3069, Ap
917	15	0.5	510	10	US-09-452-239-9	Sequence 9, Appli
918	15	0.5	510	10	US-09-974-300-809	Sequence 809, App
919	15	0.5	511	10	US-09-783-590-5135	Sequence 5135, Ap
920	15	0.5	513	10	US-09-783-590-2724	Sequence 2724, Ap
c 921	15	0.5	515	9	US-09-736-457-151	Sequence 151, App
c 922	15	0.5	515	9	US-09-902-941-151	Sequence 151, App
c 923	15	0.5	515	9	US-09-849-626-151	Sequence 151, App
924	15	0.5	517	10	US-09-998-598-1387	Sequence 1387, Ap
c 925	15	0.5	518	9	US-10-046-935-1816	Sequence 1816, Ap
c 926	15	0.5	518	9	US-09-878-178-1816	Sequence 1816, Ap
927	15	0.5	518	10	US-09-998-598-692	Sequence 692, App
c 928	15	0.5	519	10	US-09-864-761-15890	Sequence 15890, A
929	15	0.5	520	9	US-09-933-797-8	Sequence 8, Appli
c 930	15	0.5	524	10	US-09-864-761-15293	Sequence 15293, A
931	15	0.5	527	10	US-09-925-297-272	Sequence 272, App
c 932	15	0.5	529	9	US-09-854-133-682	Sequence 682, App
933	15	0.5	530	9	US-10-073-353-4	Sequence 4, Appli
c 934	15	0.5	530	10	US-09-867-550-1383	Sequence 1383, Ap
935	15	0.5	531	9	US-09-894-844-91	Sequence 91, Appl
936	15	0.5	542	9	US-09-796-692-4370	Sequence 4370, Ap
937	15	0.5	542	10	US-09-864-761-8238	Sequence 8238, Ap
c 938	15	0.5	542	10	US-09-878-574-4231	Sequence 4231, Ap
c 939	15	0.5	543	10	US-09-887-576-786	Sequence 786, App
940	15	0.5	543	10	US-09-974-300-5138	Sequence 5138, Ap
941	15	0.5	545	10	US-09-925-299-382	Sequence 382, App
942	15	0.5	546	10	US-09-815-242-4036	Sequence 4036, Ap
c 943	15	0.5	550	10	US-09-815-343-1010	Sequence 1010, Ap
944	15	0.5	552	9	US-09-910-664-10	Sequence 10, Appl
c 945	15	0.5	554	10	US-09-917-800A-1602	Sequence 1602, Ap
946	15	0.5	556	10	US-09-828-644-7	Sequence 7, Appli
c 947	15	0.5	557	9	US-09-774-639-66	Sequence 66, Appl
948	15	0.5	559	10	US-09-810-264-42	Sequence 42, Appl
c 949	15	0.5	560	9	US-10-046-935-208	Sequence 208, App
c 950	15	0.5	560	9	US-09-878-178-208	Sequence 208, App
951	15	0.5	560	10	US-09-867-701-5942	Sequence 5942, Ap
952	15	0.5	562	9	US-09-854-133-463	Sequence 463, App
c 953	15	0.5	562	10	US-09-864-761-9787	Sequence 9787, Ap
c 954	15	0.5	562	10	US-09-864-761-16224	Sequence 16224, A
955	15	0.5	562	10	US-09-738-973-463	Sequence 463, App
956	15	0.5	563	12	US-10-062-254-149	Sequence 149, App
c 957	15	0.5	566	10	US-09-925-300-867	Sequence 867, App
958	15	0.5	568	10	US-09-864-761-12790	Sequence 12790, A
c 959	15	0.5	572	10	US-09-964-238-12	Sequence 12, Appl
960	15	0.5	573	9	US-09-796-692-8943	Sequence 8943, Ap
961	15	0.5	573	10	US-09-878-574-4371	Sequence 4371, Ap
962	15	0.5	576	9	US-09-738-626-3297	Sequence 3297, Ap
963	15	0.5	577	10	US-09-864-761-7498	Sequence 7498, Ap
c 964	15	0.5	582	10	US-09-764-847-47	Sequence 47, Appl
965	15	0.5	584	10	US-09-995-494-36	Sequence 36, Appl
966	15	0.5	586	10	US-09-864-761-9940	Sequence 9940, Ap
967	15	0.5	589	10	US-09-917-800A-332	Sequence 332, App
968	15	0.5	594	9	US-09-933-797-745	Sequence 745, App

```
c 969    15   0.5   597  10  US-09-912-020-200     Sequence 200, App
c 970    15   0.5   600  10  US-09-770-149-972     Sequence 972, App
  971    15   0.5   603  9   US-09-764-904-72      Sequence 72, Appl
  972    15   0.5   603  10  US-09-764-860-600     Sequence 600, App
c 973    15   0.5   605  9   US-10-046-935-1139    Sequence 1139, Ap
c 974    15   0.5   605  9   US-09-878-178-1139    Sequence 1139, Ap
  975    15   0.5   608  10  US-09-813-358-167     Sequence 167, App
c 976    15   0.5   621  9   US-09-950-933A-5      Sequence 5, Appli
  977    15   0.5   621  10  US-09-912-020-134     Sequence 134, App
c 978    15   0.5   625  9   US-09-854-133-628     Sequence 628, App
c 979    15   0.5   625  10  US-09-833-790-298     Sequence 298, App
c 980    15   0.5   631  10  US-09-770-149-684     Sequence 684, App
c 981    15   0.5   632  10  US-09-770-149-678     Sequence 678, App
c 982    15   0.5   637  10  US-09-878-574-4574    Sequence 4574, Ap
  983    15   0.5   645  12  US-10-036-869-22      Sequence 22, Appl
c 984    15   0.5   650  12  US-10-078-929-185     Sequence 185, App
  985    15   0.5   654  10  US-09-995-598-19      Sequence 19, Appl
c 986    15   0.5   660  9   US-10-055-364-47      Sequence 47, Appl
  987    15   0.5   660  9   US-09-950-933A-18     Sequence 18, Appl
  988    15   0.5   660  10  US-09-764-847-1674    Sequence 1674, Ap
  989    15   0.5   672  9   US-09-854-133-462     Sequence 462, App
  990    15   0.5   672  10  US-09-738-973-462     Sequence 462, App
c 991    15   0.5   687  10  US-09-778-320-218     Sequence 218, App
c 992    15   0.5   687  10  US-09-780-717-30      Sequence 30, Appl
  993    15   0.5   687  10  US-09-770-149-313     Sequence 313, App
c 994    15   0.5   687  10  US-09-910-689-218     Sequence 218, App
c 995    15   0.5   687  10  US-09-765-231A-37     Sequence 37, Appl
c 996    15   0.5   687  12  US-10-010-742-218     Sequence 218, App
c 997    15   0.5   688  9   US-09-764-868-218     Sequence 218, App
  998    15   0.5   693  10  US-09-974-300-1576    Sequence 1576, Ap
  999    15   0.5   695  10  US-09-910-943-291     Sequence 291, App
 1000    15   0.5   707  10  US-09-917-800A-1530   Sequence 1530, Ap
```

ALIGNMENTS

```
RESULT 1
US-09-928-457-95
; Sequence 95, Application US/09928457
; Patent No. US20020164603A1
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA, specific proteins and peptides
;    TITLE OF INVENTION:  of the Neisseria meningitidis species bacteria, method
;    TITLE OF INVENTION:  for obtaining them and their biological application.
;    NUMBER OF SEQUENCES:  99
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (OEB)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/928,457
;      FILING DATE:  2001-08-14
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/214,759
;      FILING DATE:  199-12-10
;  INFORMATION FOR SEQ ID NO:  95:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  286 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-928-457-95

  Query Match             5.5%;  Score 177;  DB 9;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-83;
  Matches  177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy    925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy    985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 1041
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 177


RESULT 2
US-09-867-701-2325
; Sequence 2325, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2325
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(284)
;   OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2325

  Query Match             0.6%;  Score 19;  DB 10;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 8;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1263 GGCTCTGCTGCTGCAGAAA 1281
          |||||||||||||||||||
Db    185 GGCTCTGCTGCTGCAGAAA 203


RESULT 3
US-09-925-299-125
; Sequence 125, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
;   LENGTH: 1267
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-299-125

  Query Match             0.6%;  Score 19;  DB 10;  Length 1267;
  Best Local Similarity   100.0%;  Pred. No. 8.5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy     86 TAGGCGGCGGCGGCGGCAC 104
          |||||||||||||||||||
Db    588 TAGGCGGCGGCGGCGGCAC 606


RESULT 4
US-09-880-107-2355/c
; Sequence 2355, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2355
;   LENGTH: 1551
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 M55671
US-09-880-107-2355

  Query Match             0.6%;  Score 19;  DB 10;  Length 1551;
  Best Local Similarity   100.0%;  Pred. No. 8.5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1258 ACGGCGGCTCTGCTGCTGC 1276
          |||||||||||||||||||
Db   1430 ACGGCGGCTCTGCTGCTGC 1412


RESULT 5
US-09-925-301-306
; Sequence 306, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 306
;   LENGTH: 1987
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (731)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-306

  Query Match             0.6%;  Score 19;  DB 10;  Length 1987;
  Best Local Similarity   100.0%;  Pred. No. 8.6;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1263 GGCTCTGCTGCTGCAGAAA 1281
          |||||||||||||||||||
Db   1235 GGCTCTGCTGCTGCAGAAA 1253


RESULT 6
US-09-974-300-4730
; Sequence 4730, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4730
;   LENGTH: 213
;   TYPE: DNA
;   ORGANISM: Bacillus clausii
US-09-974-300-4730

  Query Match             0.6%;  Score 18;  DB 10;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1179 TGAAGCAAGCGTCCGTTT 1196
          ||||||||||||||||||
Db    163 TGAAGCAAGCGTCCGTTT 180


RESULT 7
US-09-878-574-15701
; Sequence 15701, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 15701
;   LENGTH: 276
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701070414H1
US-09-878-574-15701

  Query Match             0.6%;  Score 18;  DB 10;  Length 276;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    164 CAGCGAAATCAGCAGCAG 181
          ||||||||||||||||||
Db     85 CAGCGAAATCAGCAGCAG 102


RESULT 8
US-09-728-445-822/c
; Sequence 822, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
;  TITLE OF INVENTION:  Animals
;  FILE REFERENCE: LEX-0102-USA
;  CURRENT APPLICATION NUMBER: US/09/728,445
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/168,358
;  PRIOR FILING DATE: 1999-12-01
;  NUMBER OF SEQ ID NOS: 891
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 822
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(495)
;   OTHER INFORMATION: n = A,T,C or G
US-09-728-445-822

  Query Match             0.6%;  Score 18;  DB 10;  Length 495;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    427 GGCAGCATATCCTTTCCC 444
          ||||||||||||||||||
Db     88 GGCAGCATATCCTTTCCC 71


RESULT 9
US-09-974-300-2315/c
; Sequence 2315, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2315
;   LENGTH: 1044
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-2315

  Query Match             0.6%;  Score 18;  DB 10;  Length 1044;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    845 CCGACCATTTCCAAATAG 862
          ||||||||||||||||||
Db    205 CCGACCATTTCCAAATAG 188


RESULT 10
US-10-098-841-169
; Sequence 169, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2
;  CURRENT APPLICATION NUMBER: US/10/098,841
;  CURRENT FILING DATE:  2002-03-13
;  PRIOR APPLICATION NUMBER: 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 331
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 169
;   LENGTH: 1169
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (114)..(887)
US-10-098-841-169

  Query Match             0.6%;  Score 18;  DB 9;  Length 1169;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1063 GAAAAAGATGCTCAAAAA 1080
          ||||||||||||||||||
Db    602 GAAAAAGATGCTCAAAAA 619


RESULT 11
US-10-071-766-34
; Sequence 34, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
;  APPLICANT: Huei-Mei Chen
;  TITLE OF INVENTION: GENES EXPRESSED IN SENESCENCE
;  FILE REFERENCE: PA-0043 US
;  CURRENT APPLICATION NUMBER: US/10/071,766
;  CURRENT FILING DATE:  2002-02-07
;  NUMBER OF SEQ ID NOS: 144
;  SOFTWARE: PERL Program
; SEQ ID NO 34
;   LENGTH: 2321
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020192678A1 1382924.50
US-10-071-766-34

  Query Match             0.6%;  Score 18;  DB 9;  Length 2321;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1063 GAAAAAGATGCTCAAAAA 1080
          ||||||||||||||||||
Db   2294 GAAAAAGATGCTCAAAAA 2311
```

```
RESULT 12
US-09-815-242-7983
; Sequence 7983, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7983
;   LENGTH: 4446
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(4446)
US-09-815-242-7983

  Query Match             0.6%;  Score 18;  DB 10;  Length 4446;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1819 ATGACCGGCGGCAAGCTG 1836
          ||||||||||||||||||
Db   3904 ATGACCGGCGGCAAGCTG 3921


RESULT 13
US-09-070-927A-238/c
; Sequence 238, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/070,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/066,009
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB369
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 238:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 29729 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-09-070-927A-238

  Query Match             0.6%;  Score 18;  DB 10;  Length 29729;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1064 AAAAAGATGCTCAAAAAG 1081
          ||||||||||||||||||
Db   9605 AAAAAGATGCTCAAAAAG 9588


RESULT 14
US-09-789-836-22
; Sequence 22, Application US/09789836
; Patent No. US20020082204A1
; GENERAL INFORMATION:
;  APPLICANT: BRIGHAM, KENNETH L.
;  APPLICANT:  STECENKO, ARLENE A.
;  APPLICANT:  SEALY, LINDA
;  TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20
;  FILE REFERENCE: N-6977
;  CURRENT APPLICATION NUMBER: US/09/789,836
;  CURRENT FILING DATE:  2001-02-20
;  PRIOR APPLICATION NUMBER: 60/183,584
;  PRIOR FILING DATE: 2000-02-18
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
;   LENGTH: 102
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Oligonucleotide
US-09-789-836-22

  Query Match             0.5%;  Score 17;  DB 10;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     87 AGGCGGCGGCGGCGGCA 103
          |||||||||||||||||
Db     23 AGGCGGCGGCGGCGGCA 39


RESULT 15
US-09-789-836-21/c
```

```
; Sequence 21, Application US/09789836
; Patent No. US20020082204A1
; GENERAL INFORMATION:
;  APPLICANT: BRIGHAM, KENNETH L.
;  APPLICANT:  STECENKO, ARLENE A.
;  APPLICANT:  SEALY, LINDA
;  TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20
;  FILE REFERENCE: N-6977
;  CURRENT APPLICATION NUMBER: US/09/789,836
;  CURRENT FILING DATE:  2001-02-20
;  PRIOR APPLICATION NUMBER: 60/183,584
;  PRIOR FILING DATE: 2000-02-18
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 105
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Oligonucleotide
US-09-789-836-21

  Query Match              0.5%;  Score 17;  DB 10;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db      83 AGGCGGCGGCGGCGGCA 67


RESULT 16
US-09-789-836-33/c
; Sequence 33, Application US/09789836
; Patent No. US20020082204A1
; GENERAL INFORMATION:
;  APPLICANT: BRIGHAM, KENNETH L.
;  APPLICANT:  STECENKO, ARLENE A.
;  APPLICANT:  SEALY, LINDA
;  TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20
;  FILE REFERENCE: N-6977
;  CURRENT APPLICATION NUMBER: US/09/789,836
;  CURRENT FILING DATE:  2001-02-20
;  PRIOR APPLICATION NUMBER: 60/183,584
;  PRIOR FILING DATE: 2000-02-18
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
;   LENGTH: 106
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Oligonucleotide
US-09-789-836-33

  Query Match              0.5%;  Score 17;  DB 10;  Length 106;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      87 AGGCGGCGGCGGCGGCA 103
           |||||||||||||||||
Db      84 AGGCGGCGGCGGCGGCA 68


RESULT 17
US-09-783-590-5143
; Sequence 5143, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5143
;   LENGTH: 122
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (15)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (60)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (109)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (120)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5143

  Query Match              0.5%;  Score 17;  DB 10;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      88 GGCGGCGGCGGCGGCAC 104
           |||||||||||||||||
Db      35 GGCGGCGGCGGCGGCAC 51


RESULT 18
US-09-974-300-4131/c
; Sequence 4131, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4131
;   LENGTH: 126
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-4131

  Query Match              0.5%;  Score 17;  DB 10;  Length 126;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1510 AAAAAAGGCGGCAGCCA 1526
           |||||||||||||||||
Db      94 AAAAAAGGCGGCAGCCA 78
```

```
RESULT 19
US-09-878-574-8654
; Sequence 8654, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8654
;   LENGTH: 224
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701101486H1
US-09-878-574-8654

  Query Match             0.5%;  Score 17;  DB 10;  Length 224;
  Best Local Similarity   100.0%;  Pred. No. 90;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1353 CGTGGACAGCAAGTTCG 1369
          |||||||||||||||||
Db    172 CGTGGACAGCAAGTTCG 188


RESULT 20
US-09-880-107-3632
; Sequence 3632, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3632
;   LENGTH: 335
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 W87606
US-09-880-107-3632

  Query Match             0.5%;  Score 17;  DB 10;  Length 335;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1258 ACGGCGGCTCTGCTGCT 1274
          |||||||||||||||||
Db    120 ACGGCGGCTCTGCTGCT 136


RESULT 21
US-09-974-300-7591/c
; Sequence 7591, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7591
;   LENGTH: 336
;   TYPE: DNA
;   ORGANISM: Bacillus clausii
US-09-974-300-7591

  Query Match             0.5%;  Score 17;  DB 10;  Length 336;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2768 ACGTTCCGTTTGCCGCA 2784
          |||||||||||||||||
Db    115 ACGTTCCGTTTGCCGCA 99


RESULT 22
US-09-983-965-4734/c
; Sequence 4734, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 4734
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 17-LIB34-006-Q1-E1-E9
US-09-983-965-4734

  Query Match             0.5%;  Score 17;  DB 10;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    524 AAGACGGAGGCGGTAAA 540
          |||||||||||||||||
Db     70 AAGACGGAGGCGGTAAA 54


RESULT 23
US-10-040-739-1122/c
; Sequence 1122, Application US/10040739
; Patent No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
```

```
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;         TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;         NUMBER OF SEQUENCES: 1519
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Genetics Institute, Inc.
;              STREET: 87 CambridgePark Drive
;              CITY: Cambridge
;              STATE: Massachusetts
;              COUNTRY: U.S.A
;              ZIP: 02140
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy Disk
;              COMPUTER: IBM PC Compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/040,739
;              FILING DATE: 07-Jan-2002
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/036,520
;              FILING DATE: 03-JUN-1998
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Brown, Scott A.
;              REGISTRATION NUMBER: 32,724
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617) 498-8224
;              TELEFAX: (617) 876-5851
;    INFORMATION FOR SEQ ID NO: 1122:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 388 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: double
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         SEQUENCE DESCRIPTION: SEQ ID NO: 1122:
US-10-040-739-1122

  Query Match              0.5%;  Score 17;  DB 9;  Length 388;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    741 CGAAATAATGTCTGCAG 757
          |||||||||||||||||
Db    275 CGAAATAATGTCTGCAG 259


RESULT 24
US-09-983-965-5116
; Sequence 5116, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 5116
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 33-LIB34-008-Q1-E1-A2
US-09-983-965-5116

  Query Match              0.5%;  Score 17;  DB 10;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    156 AGCAACAACAGCGAAAT 172
          |||||||||||||||||
Db    398 AGCAACAACAGCGAAAT 414


RESULT 25
US-09-854-133-339/c
; Sequence 339, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 339
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(440)
;   OTHER INFORMATION: n = A,T,C or G
US-09-854-133-339

  Query Match              0.5%;  Score 17;  DB 9;  Length 440;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     87 AGGCGGCGGCGGCGGCA 103
          |||||||||||||||||
Db     59 AGGCGGCGGCGGCGGCA 43


RESULT 26
US-09-738-973-339/c
; Sequence 339, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Elliot, Mark
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Kalos, Michael D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.475C9
;  CURRENT APPLICATION NUMBER: US/09/738,973
;  CURRENT FILING DATE:  2000-12-14
;  NUMBER OF SEQ ID NOS:  587
;  SOFTWARE:  FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 339
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(440)
;   OTHER INFORMATION: n = A,T,C or G
US-09-738-973-339

  Query Match              0.5%;  Score 17;  DB 10;  Length 440;
  Best Local Similarity  100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       87 AGGCGGCGGCGGCGGCA 103
            |||||||||||||||||
Db       59 AGGCGGCGGCGGCGGCA 43


RESULT 27
US-09-815-242-7839
; Sequence 7839, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7839
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(441)
US-09-815-242-7839

  Query Match              0.5%;  Score 17;  DB 10;  Length 441;
  Best Local Similarity  100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1524 CCAACTGCAACTGCACG 1540
            |||||||||||||||||
Db      315 CCAACTGCAACTGCACG 331


RESULT 28
US-09-864-761-204/c
; Sequence 204, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 204
;   LENGTH: 480
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL049780.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
US-09-864-761-204

  Query Match              0.5%;  Score 17;  DB 10;  Length 480;
  Best Local Similarity  100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     2595 CATTAGTCTGTTTGCAG 2611
```

```
                ||||||||||||||||
Db      278 CATTAGTCTGTTTGCAG 262


RESULT 29
US-10-040-739-17/c
; Sequence 17, Application US/10040739
; Patent No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;        NUMBER OF SEQUENCES: 1519
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy Disk
;             COMPUTER: IBM PC Compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,739
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/036,520
;             FILING DATE: 03-JUN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 17:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 484 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-040-739-17

  Query Match             0.5%;  Score 17;  DB 9;  Length 484;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      741 CGAAATAATGTCTGCAG 757
            ||||||||||||||||
Db      299 CGAAATAATGTCTGCAG 283


RESULT 30
US-09-881-752A-225/c
; Sequence 225, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
;  APPLICANT: Kleanthous, Harold
;  APPLICANT:  Al-Garawi, Amal
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Tomb, Jean-Francois
;  APPLICANT:  Oomen, Raymond P.
;  TITLE OF INVENTION: Identification of Polynucleotides
;  TITLE OF INVENTION:  Encoding No. US20020115078A1el Helicobacter Polypeptides in t
;  TITLE OF INVENTION:  Genome
;  FILE REFERENCE: 06132/041002
;  CURRENT APPLICATION NUMBER: US/09/881,752A
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: US 08/833,457
;  PRIOR FILING DATE: 1997-04-01
;  NUMBER OF SEQ ID NOS: 370
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 225
;   LENGTH: 490
;   TYPE: DNA
;   ORGANISM: Helicobacter pylori
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (51)...(437)
US-09-881-752A-225

  Query Match             0.5%;  Score 17;  DB 10;  Length 490;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      991 ATGCTTTTCATTTTTTC 1007
            ||||||||||||||||
Db       61 ATGCTTTTCATTTTTTC 45


RESULT 31
US-09-864-761-8567/c
; Sequence 8567, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
```

```
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8567
;   LENGTH: 532
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL049780.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
US-09-864-761-8567

  Query Match             0.5%;  Score 17;  DB 10;  Length 532;
  Best Local Similarity   100.0%;  Pred. No. 93;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2595 CATTAGTCTGTTTGCAG 2611
          |||||||||||||||||
Db     43 CATTAGTCTGTTTGCAG 27


RESULT 32
US-10-108-915-31/c
; Sequence 31, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT:  Shen, Jennie
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
;  FILE REFERENCE: BB1286 US NA
;  CURRENT APPLICATION NUMBER: US/10/108,915
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
;   LENGTH: 556
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (424)
;   NAME/KEY: unsure
;   LOCATION: (499)
;   NAME/KEY: unsure
;   LOCATION: (535)
;   NAME/KEY: unsure
;   LOCATION: (554)
US-10-108-915-31

  Query Match             0.5%;  Score 17;  DB 9;  Length 556;
  Best Local Similarity   100.0%;  Pred. No. 93;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     87 AGGCGGCGGCGGCGGCA 103
          |||||||||||||||||
Db    157 AGGCGGCGGCGGCGGCA 141


RESULT 33
US-09-864-761-12180/c
; Sequence 12180, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12180
;   LENGTH: 578
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC022020.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.58
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.51
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.49
US-09-864-761-12180

  Query Match             0.5%;  Score 17;  DB 10;  Length 578;
  Best Local Similarity   100.0%;  Pred. No. 93;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    576 AGAGACTGAAGCAAAGC 592
          |||||||||||||||||
Db     27 AGAGACTGAAGCAAAGC 11


RESULT 34
```

```
US-09-738-626-2896
; Sequence 2896, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2896
;   LENGTH: 747
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-2896

  Query Match             0.5%;  Score 17;  DB 9;  Length 747;
  Best Local Similarity   100.0%;  Pred. No. 94;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2401 GGTCTGCGCGTCATCGC 2417
          |||||||||||||||||
Db     37 GGTCTGCGCGTCATCGC 53


RESULT 35
US-09-764-864-326/c
; Sequence 326, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 326
;   LENGTH: 872
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-864-326

  Query Match             0.5%;  Score 17;  DB 10;  Length 872;
  Best Local Similarity   100.0%;  Pred. No. 94;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     88 GGCGGCGGCGGCGGCAC 104
          |||||||||||||||||
Db    105 GGCGGCGGCGGCGGCAC 89


RESULT 36
US-09-938-842A-2261
; Sequence 2261, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2261
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2261

  Query Match             0.5%;  Score 17;  DB 9;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2456 GTGTTGAAGGCAAAATG 2472
          |||||||||||||||||
Db    317 GTGTTGAAGGCAAAATG 333


RESULT 37
US-09-974-300-5167
; Sequence 5167, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5167
;   LENGTH: 960
;   TYPE: DNA
;   ORGANISM: Bacillus clausii
US-09-974-300-5167

  Query Match             0.5%;  Score 17;  DB 10;  Length 960;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2944 AAAGCCGTCCTGTTTGC 2960
          |||||||||||||||||
Db    333 AAAGCCGTCCTGTTTGC 349


RESULT 38
US-09-789-836-2/c
; Sequence 2, Application US/09789836
; Patent No. US20020082204A1
; GENERAL INFORMATION:
;  APPLICANT: BRIGHAM, KENNETH L.
;  APPLICANT:  STECENKO, ARLENE A.
;  APPLICANT:  SEALY, LINDA
```

```
;  TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20
;  FILE REFERENCE: N-6977
;  CURRENT APPLICATION NUMBER: US/09/789,836
;  CURRENT FILING DATE:  2001-02-20
;  PRIOR APPLICATION NUMBER: 60/183,584
;  PRIOR FILING DATE: 2000-02-18
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1038
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-789-836-2

  Query Match              0.5%;  Score 17;  DB 10;  Length 1038;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       87 AGGCGGCGGCGGCGGCA 103
            |||||||||||||||||
Db       57 AGGCGGCGGCGGCGGCA 41


RESULT 39
US-09-822-830A-366/c
; Sequence 366, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
;  APPLICANT: Genetics Institute, Inc.
;  APPLICANT:  Wong, Gordon G.
;  APPLICANT:  Clark, Hilary
;  APPLICANT:  Fechtel, Kim
;  APPLICANT:  Agostino, Michael J.
;  APPLICANT:  Howes, Steven H.
;  APPLICANT:  Resnick, Richard J.
;  APPLICANT:  Gulukota, Kamalakar
;  APPLICANT:  Graham, James R.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
;  FILE REFERENCE: GIN 6402
;  CURRENT APPLICATION NUMBER: US/09/822,830A
;  CURRENT FILING DATE:  2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195,604
;  PRIOR FILING DATE: 2000-04-06
;  NUMBER OF SEQ ID NOS: 631
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 366
;   LENGTH: 1408
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1254,1257
;   OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-366

  Query Match              0.5%;  Score 17;  DB 10;  Length 1408;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     2748 GGGCGCACTGGGCGGTG 2764
            |||||||||||||||||
Db      922 GGGCGCACTGGGCGGTG 906


RESULT 40
US-10-108-915-33/c
; Sequence 33, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT:  Shen, Jennie
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
;  FILE REFERENCE: BB1286 US NA
;  CURRENT APPLICATION NUMBER: US/10/108,915
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
;   LENGTH: 1428
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-10-108-915-33

  Query Match              0.5%;  Score 17;  DB 9;  Length 1428;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       87 AGGCGGCGGCGGCGGCA 103
            |||||||||||||||||
Db      169 AGGCGGCGGCGGCGGCA 153


Search completed: January 28, 2003, 00:31:35
Job time : 250 secs
```

102(a)
102(e)

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 19:09:24 ; Search time 433 Seconds
(without alignments)
16663.747 Million cell updates/sec

Title: US-09-830-433A-7
Perfect score: 3204
Sequence: 1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

NO

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3204	100.0	3204	21	AAA15299	DNA encoding a pol
	2	990	30.9	1365	21	AAZ53301	Neisseria meningit
	3	834	26.0	1365	21	AAZ53300	Neisseria meningit
	4	834	26.0	4218	22	AAS43905	Neisseria meningit
	5	834	26.0	4218	22	AAD17058	N. meningitidis st
c	6	834	26.0	13423	21	AAA81527	N. meningitidis pa
	7	834	26.0	349980	21	AAF21612	Neisseria meningit
	8	831	25.9	3939	22	AAS43875	Neisseria meningit
	9	831	25.9	3939	22	AAS43880	Neisseria meningit
	10	831	25.9	3939	22	AAD17040	N. meningitidis st
	11	831	25.9	3939	22	AAD17045	N. meningitidis st
	12	831	25.9	4170	22	AAS43902	Neisseria meningit
	13	831	25.9	4170	22	AAD17055	N. meningitidis st
	14	831	25.9	4179	22	AAS43877	Neisseria meningit
	15	831	25.9	4179	22	AAD17042	N. meningitidis st
	16	831	25.9	4335	22	AAS43899	Neisseria meningit
	17	831	25.9	4335	22	AAD17052	N. meningitidis st
	18	831	25.9	4344	22	AAS43876	Neisseria meningit
	19	831	25.9	4425	22	AAS43874	Neisseria meningit
	20	831	25.9	4425	22	AAD17039	N. meningitidis st
	21	398	12.4	3798	22	AAD17041	N. meningitidis st
	22	189	5.9	526	21	AAZ53298	Neisseria meningit
	23	177	5.5	286	19	AAV03603	Neisseria meningit
	24	177	5.5	286	21	AAA15396	Genomic fragment o
	25	160	5.0	528	21	AAZ53297	Neisseria meningit
	26	149	4.7	537	21	AAZ53296	Neisseria gonorrhe
	27	131	4.1	1365	21	AAZ53299	Neisseria gonorrhe
c	28	57	1.8	476	21	AAA81823	N. meningitidis pa
	29	41	1.3	577	21	AAA82003	N. meningitidis pa
	30	35	1.1	50	22	AAS44248	Neisseria meningit
	31	25	0.8	2280	23	ABL05527	Drosophila melanog
c	32	25	0.8	3302	23	ABL28074	Drosophila melanog
c	33	25	0.8	13327	23	ABL05526	Drosophila melanog
c	34	23	0.7	30	21	AAA15341	PCR primer for a N
	35	22	0.7	2382	23	ABL27930	Drosophila melanog
	36	21	0.7	34	21	AAA15340	PCR primer for a N
	37	20	0.6	34	21	AAZ55456	Neisseria species
c	38	20	0.6	735	18	AAT45009	Serratia marcescen
c	39	20	0.6	735	19	AAV43062	DNA encoding an ac
c	40	20	0.6	1105	24	ABQ32992	Oligonucleotide fo
	41	20	0.6	1105	24	ABQ32993	Oligonucleotide fo
c	42	20	0.6	5019	22	AAF82427	Moraxella catarrha
	43	20	0.6	15004	23	ABL27628	Drosophila melanog
c	44	20	0.6	89047	22	AAF28547	Genomic fragment #
	45	19	0.6	29	22	AAS44242	Neisseria meningit
	46	19	0.6	32	22	AAS44246	Neisseria meningit
	47	19	0.6	284	24	ABL79347	Human ovarian canc
c	48	19	0.6	455	23	ABV14706	Human prostate exp
c	49	19	0.6	570	22	AAH12092	Human cDNA clone (
	50	19	0.6	807	22	AAH52244	Human AFP protein
	51	19	0.6	1267	21	AAC98115	Human colon cancer
	52	19	0.6	1286	22	AAI93900	Human stomach canc
	53	19	0.6	1286	22	AAH16346	Human cDNA sequenc
	54	19	0.6	1383	22	AAH52253	Human AFP protein
	55	19	0.6	1451	22	AAH13753	Human cDNA sequenc
c	56	19	0.6	1551	24	ABN95857	Gene #2355 used to
	57	19	0.6	1596	22	AAH52254	Human AFP protein
	58	19	0.6	1611	22	AAK87974	Human digestive sy
	59	19	0.6	1987	21	AAC77912	Human cancer assoc
	60	19	0.6	2000	13	AAQ25783	CTP synthase gene
	61	19	0.6	3001	12	AAQ14668	CTP synthetase gen
	62	19	0.6	3017	24	ABL55297	Human mitochondria
	63	19	0.6	4424	22	AAK51737	Human polynucleoti
	64	19	0.6	4504	22	AAK52721	Human polynucleoti
c	65	19	0.6	11340	23	ABL16852	Drosophila melanog
	66	19	0.6	17687	22	AAS42069	Genomic sequence #
	67	19	0.6	19815	22	AAS42064	Genomic sequence #
	68	19	0.6	62909	22	AAF28545	Genomic fragment #
c	69	18	0.6	28	22	AAS44225	Neisseria meningit
c	70	18	0.6	28	22	AAD17068	N. meningitidis fu
	71	18	0.6	30	22	AAD15111	Extended_TOGA prim
c	72	18	0.6	60	16	AAT00248	Thrombin 60N DNA l
c	73	18	0.6	60	22	AAF70800	Thrombin high affi
	74	18	0.6	111	22	AAF26369	P. putida oxygenas
	75	18	0.6	213	24	ABK77439	Bacillus clausii g
c	76	18	0.6	267	20	AAV90490	EST clone DN714.
c	77	18	0.6	267	20	AAV89054	EST clone CB123.
	78	18	0.6	288	21	AAC76303	Human ORFX ORF1858
c	79	18	0.6	378	16	AAT24092	Human gene signatu
c	80	18	0.6	504	23	ABV48468	Human prostate exp
	81	18	0.6	507	24	ABQ30826	Oligonucleotide fo
c	82	18	0.6	507	24	ABQ30827	Oligonucleotide fo

c	83	18	0.6	592	22	AAH09157	Human cDNA clone (
	84	18	0.6	593	24	ABQ51880	Oligonucleotide fo
c	85	18	0.6	593	24	ABQ51881	Oligonucleotide fo
	86	18	0.6	633	23	AAS67168	DNA encoding novel
c	87	18	0.6	653	22	AAH34289	Human colon cancer
	88	18	0.6	720	24	ABQ33460	Oligonucleotide fo
c	89	18	0.6	720	24	ABQ33461	Oligonucleotide fo
	90	18	0.6	797	24	ABQ46156	Oligonucleotide fo
c	91	18	0.6	797	24	ABQ46157	Oligonucleotide fo
	92	18	0.6	885	24	AAD24396	Human RNA metaboli
	93	18	0.6	888	22	AAH24356	Human ribosome sub
c	94	18	0.6	962	21	AAC93388	Human secreted pro
	95	18	0.6	994	20	AAV84522	Human secreted pro
	96	18	0.6	994	22	ABA83305	Human secreted pro
c	97	18	0.6	1044	24	ABK75024	Bacillus lichenifo
c	98	18	0.6	1104	23	ABL24461	Drosophila melanog
	99	18	0.6	1136	22	AAI59751	Human polynucleoti
	100	18	0.6	1139	21	AAC76370	Human ORFX ORF1925
	101	18	0.6	1169	22	AAI57965	Human polynucleoti
	102	18	0.6	1177	22	AAF32742	Human secreted pro
	103	18	0.6	1183	24	AAL40191	Isoprenoid related
c	104	18	0.6	1335	23	ABL22177	Drosophila melanog
c	105	18	0.6	1375	21	AAZ33365	Human secreted pro
	106	18	0.6	1431	24	ABQ42300	Oligonucleotide fo
c	107	18	0.6	1431	24	ABQ42301	Oligonucleotide fo
c	108	18	0.6	1434	23	AAS80223	DNA encoding novel
	109	18	0.6	1501	24	ABQ43092	Oligonucleotide fo
c	110	18	0.6	1501	24	ABQ43093	Oligonucleotide fo
	111	18	0.6	1531	23	AAS84360	DNA encoding novel
	112	18	0.6	1614	22	AAH17079	Human cDNA sequenc
c	113	18	0.6	1689	22	AAF80591	Receptor #79 parti
	114	18	0.6	1721	21	AAC79728	Human secreted pro
	115	18	0.6	1903	19	AAX14490	H. pylori GHPO 113
c	116	18	0.6	2363	20	AAX78216	Melon MEERS cDNA.
	117	18	0.6	2464	22	AAS45213	cDNA encoding nove
	118	18	0.6	2536	24	ABK35738	cDNA sequence #129
c	119	18	0.6	2682	22	AAH14265	Human cDNA sequenc
c	120	18	0.6	2726	23	ABL09441	Drosophila melanog
	121	18	0.6	2902	22	AAC90433	Murine Lymph node
	122	18	0.6	2905	23	AAS78259	DNA encoding novel
	123	18	0.6	2943	23	AAS86435	DNA encoding novel
c	124	18	0.6	3030	23	ABL17899	Drosophila melanog
c	125	18	0.6	3281	23	ABL09450	Drosophila melanog
	126	18	0.6	3766	23	ABL22176	Drosophila melanog
c	127	18	0.6	3971	23	ABL17923	Drosophila melanog
	128	18	0.6	4446	23	AAS54346	Pseudomonas aerugi
c	129	18	0.6	5080	23	ABL09440	Drosophila melanog
c	130	18	0.6	5893	24	ABL32859	Human immune syste
	131	18	0.6	5952	24	AAS61403	Human gene regulat
c	132	18	0.6	6988	23	ABL29124	Drosophila melanog
	133	18	0.6	6999	23	AAS85200	DNA encoding novel
c	134	18	0.6	7167	24	ABL32400	Human immune syste
	135	18	0.6	7800	23	ABL24460	Drosophila melanog
	136	18	0.6	7875	21	AAA92470	Shewanella sp. SCR
c	137	18	0.6	9299	24	ABL33725	Human immune syste
	138	18	0.6	10337	23	ABL25704	Drosophila melanog
c	139	18	0.6	12986	22	AAS46553	Tumour suppressor
	140	18	0.6	14542	24	ABL70191	Chemically treated
	141	18	0.6	14542	24	AAS61147	Human gene regulat
	142	18	0.6	14542	24	ABK31234	Signal transductio
c	143	18	0.6	16560	23	ABL17922	Drosophila melanog
c	144	18	0.6	16759	24	ABL36297	Human lysosomal ac
c	145	18	0.6	17893	24	ABL33365	Human immune syste
c	146	18	0.6	19165	23	ABL17898	Drosophila melanog
	147	18	0.6	23821	21	AAA92469	Shewanella sp. SCR
c	148	18	0.6	29729	20	AAX13175	Enterococcus faeca
	149	18	0.6	50925	21	AAA81487	N. meningitidis pa
	150	18	0.6	349980	21	AAF21610	Neisseria meningit
	151	18	0.6	1437668	21	AAA81490	N. meningitidis B
c	152	17	0.5	20	24	AAD37756	Real-time validati
	153	17	0.5	51	22	AAL28807	Human SNP oligonuc
c	154	17	0.5	54	20	AAZ07332	Mutated hTR promot
c	155	17	0.5	54	20	AAZ00333	Mutated hTR promot

c	156	17	0.5	54	20	AAZ00332	Mutated hTR promot
	157	17	0.5	77	17	AAT09808	Tomato genomic DNA
	158	17	0.5	102	22	AAD19388	Synthetic mutant b
c	159	17	0.5	104	23	ABK51408	Human CCAAT enhanc
c	160	17	0.5	106	22	AAD19387	Synthetic mutant t
c	161	17	0.5	106	22	AAD19399	Wild type C/EBPbet
c	162	17	0.5	107	20	AAZ07325	Mutated hTR promot
c	163	17	0.5	116	22	AAS39798	Genomic sequence #
c	164	17	0.5	116	22	AAK90151	Human digestive sy
	165	17	0.5	117	18	AAT93339	EST locus ys82e12.
c	166	17	0.5	126	24	ABK76840	Bacillus lichenifo
	167	17	0.5	134	21	AAC09407	Human secreted pro
c	168	17	0.5	176	20	AAZ07323	Human telomerase R
c	169	17	0.5	176	20	AAZ07251	Human telomerase R
c	170	17	0.5	267	24	AAD37707	Mouse digital sequ
	171	17	0.5	289	24	ABI99576	Mouse ischaemic co
	172	17	0.5	315	24	ABN79072	Human ORF4019 cDNA
c	173	17	0.5	317	22	AAH69137	Human cervical can
c	174	17	0.5	320	22	AAI89249	Human polynucleoti
	175	17	0.5	322	18	AAT93340	EST locus yx44f06.
	176	17	0.5	335	24	ABN97137	Gene #3635 used to
c	177	17	0.5	336	24	ABK80300	Bacillus clausii g
c	178	17	0.5	347	22	AAK84364	Human immune/haema
	179	17	0.5	352	22	AAF65999	Novel human polynu
	180	17	0.5	356	21	AAC10429	Human secreted pro
	181	17	0.5	362	21	AAC57588	Arachidonic acid m
	182	17	0.5	362	21	AAC57589	Arachidonic acid m
	183	17	0.5	362	21	AAC57590	Arachidonic acid m
	184	17	0.5	362	21	AAC57591	Arachidonic acid m
c	185	17	0.5	363	22	AAI82582	Human polynucleoti
c	186	17	0.5	376	22	AAI83211	Human polynucleoti
	187	17	0.5	384	21	AAC74894	Human ORFX ORF449
c	188	17	0.5	388	20	AAV88644	EST clone GY307.
c	189	17	0.5	396	21	AAC04103	Human secreted pro
	190	17	0.5	400	18	AAV77284	Staphylococcus aur
	191	17	0.5	404	23	ABV03564	Human prostate exp
	192	17	0.5	419	21	AAC09275	Human secreted pro
	193	17	0.5	420	23	ABV12733	Human prostate exp
	194	17	0.5	429	21	ABQ63154	Mycobacterium tube
c	195	17	0.5	440	21	AAC79297	Human lung tumour-
c	196	17	0.5	440	23	AAD23373	Human lung tumour-
	197	17	0.5	441	23	AAS54202	Pseudomonas aerugi
	198	17	0.5	443	23	ABV18440	Human prostate exp
c	199	17	0.5	447	23	AAS81595	DNA encoding novel
	200	17	0.5	463	23	ABV33862	Human prostate exp
	201	17	0.5	463	23	ABV42747	Human prostate exp
c	202	17	0.5	480	22	ABA51921	Human foetal liver
c	203	17	0.5	480	22	ABA21738	Probe #204 for gen
c	204	17	0.5	480	22	AAK00215	Human brain expres
c	205	17	0.5	480	22	AAK25655	Human bone marrow
c	206	17	0.5	480	22	AAI10283	Probe #216 for gen
c	207	17	0.5	480	22	AAI31531	Probe #217 used to
c	208	17	0.5	480	22	AAI00220	Probe #211 used to
c	209	17	0.5	480	22	AAH46893	cDNA encoding huma
c	210	17	0.5	480	24	ABS00227	Human genome-deriv
c	211	17	0.5	484	20	AAV87539	EST clone DD426.
	212	17	0.5	487	22	AAK92117	Human cDNA 5'-end
	213	17	0.5	487	22	AAK93380	Human cDNA clone r
c	214	17	0.5	490	19	AAX14032	H. pylori GHPO 893
	215	17	0.5	490	22	AAL04542	Human reproductive
	216	17	0.5	490	23	ABL97465	Human testicular a
	217	17	0.5	494	23	ABV48225	Human prostate exp
	218	17	0.5	500	18	AAT43451	ATM gene exon 11.
	219	17	0.5	500	22	AAF57603	ATM genomic fragme
	220	17	0.5	518	22	AAK61544	Human immune/haema
c	221	17	0.5	521	21	AAA98391	Red buffalo PTEN/M
	222	17	0.5	524	24	ABQ14858	Oligonucleotide fo
c	223	17	0.5	524	24	ABQ14859	Oligonucleotide fo
c	224	17	0.5	526	23	ABV58421	Human prostate exp
c	225	17	0.5	532	22	ABA62815	Human foetal liver
c	226	17	0.5	532	22	ABA30101	Probe #8567 for ge
c	227	17	0.5	532	22	AAK11214	Human brain expres
c	228	17	0.5	532	22	AAK37015	Human bone marrow

c	229	17	0.5	532	22	AAI17860	Probe #7793 for ge
c	230	17	0.5	532	22	AAI42831	Probe #11517 used
c	231	17	0.5	532	24	ABS11017	Human genome-deriv
c	232	17	0.5	540	19	AAX14446	H. pylori GHPO 892
c	233	17	0.5	550	24	ABK11646	Human cDNA encodin
c	234	17	0.5	578	22	AAK08520	Human brain expres
c	235	17	0.5	578	22	AAK34404	Human bone marrow
c	236	17	0.5	578	22	AAI40127	Probe #8813 used t
	237	17	0.5	581	23	ABV44584	Human prostate exp
	238	17	0.5	585	23	ABL24939	Drosophila melanog
c	239	17	0.5	585	24	ABN60642	Human cancer relat
	240	17	0.5	594	24	ABQ56232	Human ovarian anti
	241	17	0.5	594	24	ABQ32006	Oligonucleotide fo
c	242	17	0.5	594	24	ABQ32007	Oligonucleotide fo
	243	17	0.5	595	23	AAS87162	DNA encoding novel
	244	17	0.5	605	22	AAK72050	Human immune/haema
c	245	17	0.5	605	22	AAH69784	Human cervical can
	246	17	0.5	607	24	ABQ20014	Oligonucleotide fo
c	247	17	0.5	607	24	ABQ20015	Oligonucleotide fo
	248	17	0.5	611	22	AAH05991	Human cDNA clone (
c	249	17	0.5	612	22	AAH68969	Human cervical can
c	250	17	0.5	640	22	AAH70632	Human cervical can
c	251	17	0.5	646	21	AAF07854	Fusarium venenatum
c	252	17	0.5	651	24	ABK36136	cDNA sequence #527
	253	17	0.5	662	24	ABQ14048	Oligonucleotide fo
c	254	17	0.5	662	24	ABQ14049	Oligonucleotide fo
c	255	17	0.5	667	22	AAH70942	Human cervical can
c	256	17	0.5	680	18	AAT58803	Human telomerase e
	257	17	0.5	688	21	AAF11899	Aspergillus oryzae
c	258	17	0.5	697	21	AAF15055	Trichoderma reesei
c	259	17	0.5	710	22	AAH70524	Human cervical can
	260	17	0.5	710	23	AAS78555	DNA encoding novel
c	261	17	0.5	713	22	AAH08611	Human cDNA clone (
c	262	17	0.5	716	24	ABQ32384	Oligonucleotide fo
	263	17	0.5	716	24	ABQ32385	Oligonucleotide fo
	264	17	0.5	724	24	ABQ49084	Oligonucleotide fo
c	265	17	0.5	724	24	ABQ49085	Oligonucleotide fo
c	266	17	0.5	741	22	AAK93489	Human cDNA clone r
	267	17	0.5	746	24	AAS61926	Porcine muscular s
	268	17	0.5	747	22	AAH67861	C glutamicum codin
	269	17	0.5	748	22	AAK67220	Human immune/haema
	270	17	0.5	748	22	AAK77725	Human immune/haema
c	271	17	0.5	762	22	AAH05372	Human cDNA clone (
	272	17	0.5	762	22	AAH07885	Human cDNA clone (
c	273	17	0.5	781	21	AAZ33349	Human secreted pro
	274	17	0.5	782	24	ABQ18794	Oligonucleotide fo
c	275	17	0.5	782	24	ABQ18795	Oligonucleotide fo
	276	17	0.5	783	24	ABQ53582	Oligonucleotide fo
c	277	17	0.5	783	24	ABQ53583	Oligonucleotide fo
	278	17	0.5	800	24	ABQ33104	Oligonucleotide fo
c	279	17	0.5	800	24	ABQ33105	Oligonucleotide fo
	280	17	0.5	800	24	ABQ36102	Oligonucleotide fo
c	281	17	0.5	800	24	ABQ36103	Oligonucleotide fo
	282	17	0.5	803	24	AAS61925	Porcine muscular s
	283	17	0.5	811	24	ABQ38800	Oligonucleotide fo
c	284	17	0.5	811	24	ABQ38801	Oligonucleotide fo
c	285	17	0.5	814	22	AAI94797	Human neuroblastom
	286	17	0.5	821	24	ABQ37728	Oligonucleotide fo
c	287	17	0.5	821	24	ABQ37729	Oligonucleotide fo
c	288	17	0.5	826	22	AAH03359	Human cDNA clone (
c	289	17	0.5	852	22	AAI94415	Human neuroblastom
c	290	17	0.5	856	22	AAH34310	Human colon cancer
	291	17	0.5	859	20	AAZ11977	Neisseria meningit
	292	17	0.5	859	21	AAA81323	N. meningitidis Me
c	293	17	0.5	867	20	AAZ07250	Human telomerase R
c	294	17	0.5	867	23	AAS80206	DNA encoding novel
	295	17	0.5	867	23	AAS93950	DNA encoding novel
	296	17	0.5	868	24	ABQ40660	Oligonucleotide fo
c	297	17	0.5	868	24	ABQ40661	Oligonucleotide fo
	298	17	0.5	869	23	ABL17159	Drosophila melanog
	299	17	0.5	871	18	AAT93326	Exon 1 of gene of
c	300	17	0.5	872	22	AAS26147	Human cDNA encodin
	301	17	0.5	890	24	ABQ43626	Oligonucleotide fo
c	302	17	0.5	890	24	ABQ43627	Oligonucleotide fo
	303	17	0.5	897	20	AAZ11988	Neisseria gonorrho
	304	17	0.5	897	24	ABN71052	Streptococcus poly
	305	17	0.5	902	24	ABQ30942	Oligonucleotide fo
c	306	17	0.5	902	24	ABQ30943	Oligonucleotide fo
	307	17	0.5	906	22	AAC91351	Human polynucleoti
c	308	17	0.5	909	21	AAC76460	Human ORFX ORF2015
	309	17	0.5	917	24	ABQ30772	Oligonucleotide fo
c	310	17	0.5	917	24	ABQ30773	Oligonucleotide fo
c	311	17	0.5	921	24	ABQ22926	Oligonucleotide fo
	312	17	0.5	921	24	ABQ22927	Oligonucleotide fo
c	313	17	0.5	924	24	ABQ44484	Oligonucleotide fo
	314	17	0.5	924	24	ABQ44485	Oligonucleotide fo
	315	17	0.5	925	21	AAA93918	Rice bZIP gene MN3
	316	17	0.5	932	24	ABQ40354	Oligonucleotide fo
c	317	17	0.5	932	24	ABQ40355	Oligonucleotide fo
c	318	17	0.5	942	23	AAS77739	DNA encoding novel
	319	17	0.5	942	24	ABN69421	Streptococcus poly
	320	17	0.5	960	24	ABK77876	Bacillus clausii g
c	321	17	0.5	980	22	AAS09471	Human DNA containi
c	322	17	0.5	981	18	AAT89242	Clone containing h
c	323	17	0.5	981	19	AAV63644	Human telomerase R
c	324	17	0.5	981	19	AAV41194	PstI fragment cont
c	325	17	0.5	981	19	AAV19481	Human telomerase h
c	326	17	0.5	981	20	AAZ23626	Human clone 28-1 c
c	327	17	0.5	981	21	AAA37564	Human telomerase n
	328	17	0.5	981	22	AAH45398	Human ADAM10 promo
c	329	17	0.5	981	23	AAS15442	PstI fragment cont
c	330	17	0.5	981	24	AAD24234	Human telomerase (
	331	17	0.5	982	24	ABQ36674	Oligonucleotide fo
c	332	17	0.5	982	24	ABQ36675	Oligonucleotide fo
	333	17	0.5	984	22	AAF27606	Mevalonate pathway
	334	17	0.5	990	22	AAI68313	Rice NIF1 related
	335	17	0.5	1004	21	AAC44468	Zea mays DNA fragm
c	336	17	0.5	1006	21	AAZ97042	Human secreted pro
	337	17	0.5	1037	23	ABV22330	Human prostate exp
	338	17	0.5	1037	23	ABV23506	Human prostate exp
	339	17	0.5	1037	23	ABV28156	Human prostate exp
	340	17	0.5	1037	23	ABV29362	Human prostate exp
c	341	17	0.5	1038	22	AAD19380	Human CCAAT/enhanc
c	342	17	0.5	1038	22	AAS14696	Human cDNA encodin
c	343	17	0.5	1038	23	ABK51402	DNA encoding human
c	344	17	0.5	1038	24	AAL44084	Human C/EBP-beta p
c	345	17	0.5	1038	24	ABK51997	DNA encoding wild-
c	346	17	0.5	1039	24	ABQ38044	Oligonucleotide fo
	347	17	0.5	1039	24	ABQ38045	Oligonucleotide fo
c	348	17	0.5	1042	23	ABK51410	Human CCAAT enhanc
c	349	17	0.5	1045	21	AAF13783	Aspergillus oryzae
	350	17	0.5	1071	22	AAH33017	Human colon cancer
	351	17	0.5	1071	24	ABL89534	Human polynucleoti
	352	17	0.5	1084	22	AAF31621	Mycobacterium tube
	353	17	0.5	1119	24	ABQ90245	M. capsulatus gene
	354	17	0.5	1140	22	AAF31648	Mycobacterium tube
c	355	17	0.5	1149	21	AAF12939	Aspergillus oryzae
	356	17	0.5	1168	24	ABQ47030	Oligonucleotide fo
c	357	17	0.5	1168	24	ABQ47031	Oligonucleotide fo
	358	17	0.5	1174	22	AAK53200	Human polynucleoti
	359	17	0.5	1195	24	ABQ31770	Oligonucleotide fo
c	360	17	0.5	1195	24	ABQ31771	Oligonucleotide fo
c	361	17	0.5	1221	22	AAC85144	Human early growth
	362	17	0.5	1235	24	ABQ32040	Oligonucleotide fo
c	363	17	0.5	1235	24	ABQ32041	Oligonucleotide fo
	364	17	0.5	1242	20	AAZ11978	Neisseria meningit
	365	17	0.5	1242	20	AAZ11979	Neisseria meningit
	366	17	0.5	1242	20	AAZ11980	Neisseria gonorrho
	367	17	0.5	1246	22	AAF71529	Corynebacterium gl
	368	17	0.5	1255	19	AAV46408	Human tumour suppr
	369	17	0.5	1257	19	AAV46393	Human tumour suppr
	370	17	0.5	1257	20	AAX57674	Human tumour suppr
	371	17	0.5	1265	22	AAF23415	Human SEC6 DNA seq
	372	17	0.5	1267	24	ABQ40672	Oligonucleotide fo
c	373	17	0.5	1267	24	ABQ40673	Oligonucleotide fo
	374	17	0.5	1281	9	AAN81257	Mutant aroA gene e

	375	17	0.5	1281	13	AAQ29507	EPSP synthase wild
	376	17	0.5	1281	13	AAQ29508	EPSP synthase muta
	377	17	0.5	1281	15	AAQ64903	EPSP synthase gene
	378	17	0.5	1281	15	AAQ64904	Mutant EPSP syntha
	379	17	0.5	1286	23	AAS64240	DNA encoding novel
c	380	17	0.5	1318	22	AAH25841	Murine SHD1 protei
c	381	17	0.5	1372	21	AAZ40202	O. sativa nicotian
	382	17	0.5	1372	24	ABQ15720	Oligonucleotide fo
c	383	17	0.5	1372	24	ABQ15721	Oligonucleotide fo
c	384	17	0.5	1375	24	ABQ25200	Oligonucleotide fo
	385	17	0.5	1375	24	ABQ25201	Oligonucleotide fo
c	386	17	0.5	1398	23	AAS87887	DNA encoding novel
c	387	17	0.5	1408	24	AAS62579	cDNA sequence #366
c	388	17	0.5	1440	20	AAV64277	Human EGR-2 cDNA.
	389	17	0.5	1460	23	ABL20291	Drosophila melanog
	390	17	0.5	1482	21	AAC44746	Zea mays DNA fragm
	391	17	0.5	1519	21	AAC76492	Human ORFX ORF2047
	392	17	0.5	1521	23	ABL17483	Drosophila melanog
	393	17	0.5	1545	23	AAS54038	Pseudomonas aerugi
	394	17	0.5	1548	22	AAI68317	Rice NIF1 related
	395	17	0.5	1574	23	AAS75663	DNA encoding novel
	396	17	0.5	1607	23	ABL20289	Drosophila melanog
	397	17	0.5	1610	24	ABK09767	Human ovarian tumo
c	398	17	0.5	1629	18	AAT66504	Gene for rice mito
	399	17	0.5	1664	24	ABA99677	Murine osteoporosi
	400	17	0.5	1676	22	ABA19238	Human nervous syst
	401	17	0.5	1683	20	AAX89293	Human regulatory p
	402	17	0.5	1701	23	ABV25058	Human prostate exp
	403	17	0.5	1713	23	AAS91801	DNA encoding novel
c	404	17	0.5	1720	22	AAH16989	Human cDNA sequenc
	405	17	0.5	1725	19	AAV45137	Homo sapiens P-TEN
c	406	17	0.5	1725	24	ABA95144	Human lysyl-oxidas
	407	17	0.5	1756	23	AAS85527	DNA encoding novel
c	408	17	0.5	1765	20	AAZ07247	Human telomerase R
c	409	17	0.5	1766	24	ABI99779	Mouse ischaemic co
	410	17	0.5	1773	18	AAT93354	Gene of IMAGE clon
c	411	17	0.5	1773	21	AAZ29548	Zea mays RecA-like
	412	17	0.5	1776	24	ABK84225	Human cDNA differe
	413	17	0.5	1776	24	AAS17569	DNA encoding human
	414	17	0.5	1784	20	AAX07339	Dual specificity p
	415	17	0.5	1791	16	AAT04154	BCG DapB gene. My
c	416	17	0.5	1796	23	ABL09605	Drosophila melanog
	417	17	0.5	1850	24	ABQ69203	Listeria monocytog
c	418	17	0.5	1910	21	AAF21187	Human low adenosin
c	419	17	0.5	1910	21	AAA35065	Human adenosine re
c	420	17	0.5	1910	22	AAD19379	Human CCAAT/enhanc
c	421	17	0.5	1910	23	ABK51401	Human CCAAT enhanc
c	422	17	0.5	1910	24	AAL44083	Human C/EBP-beta p
c	423	17	0.5	1910	24	ABK51996	DNA encoding wild-
c	424	17	0.5	1910	24	ABK84563	Human cDNA differe
c	425	17	0.5	1910	24	ABL94237	Human C/EBP beta-e
c	426	17	0.5	1913	12	AAQ15015	Nuclear factor C/E
	427	17	0.5	1962	20	AAX57672	Mouse tumour suppr
	428	17	0.5	1967	24	ABN95920	Gene #2418 used to
	429	17	0.5	1971	12	AAQ13359	Human mevalonate k
	430	17	0.5	2004	22	AAH24555	Zinc finger protei
c	431	17	0.5	2007	22	AAH16766	Human cDNA sequenc
c	432	17	0.5	2016	22	AAS26565	Human cDNA encodin
	433	17	0.5	2022	22	AAH15029	Human cDNA sequenc
c	434	17	0.5	2037	22	AAH51958	Mycobacterium tube
	435	17	0.5	2047	22	AAK52216	Human polynucleoti
	436	17	0.5	2067	24	ABL34301	Human immune syste
	437	17	0.5	2067	24	ABK28460	DNA transcription
c	438	17	0.5	2088	22	AAH14076	Human cDNA sequenc
	439	17	0.5	2139	19	AAV56670	Human ELL2 cDNA.
	440	17	0.5	2160	20	AAX57700	Mouse MMAC1 gene.
c	441	17	0.5	2171	21	AAF21188	Human low adenosin
c	442	17	0.5	2171	21	AAA35066	Human adenosine re
	443	17	0.5	2238	23	ABL28399	Drosophila melanog
	444	17	0.5	2274	22	AAD12784	Human nuclear horm
	445	17	0.5	2305	24	ABA94651	Human secreted red
c	446	17	0.5	2328	24	AAI67786	Human lysyl oxidas
c	447	17	0.5	2406	17	AAT33677	Human transcriptio
c	448	17	0.5	2408	23	ABL14691	Drosophila melanog
c	449	17	0.5	2425	17	AAT10283	Gene for RNA compo
c	450	17	0.5	2425	17	AAT11027	DNA encoding the h
c	451	17	0.5	2426	19	AAV22994	DNA containing hum
c	452	17	0.5	2463	22	AAK94536	Human full-length
c	453	17	0.5	2467	24	ABQ91979	Human NF-kB activa
	454	17	0.5	2471	21	AAA73711	Human PITSLRE prot
	455	17	0.5	2471	24	ABL51630	Human PITSLRE prot
	456	17	0.5	2493	23	AAS88122	DNA encoding novel
c	457	17	0.5	2585	23	ABL24938	Drosophila melanog
c	458	17	0.5	2588	21	AAC76756	Human ORFX ORF2311
	459	17	0.5	2595	22	AAH33385	Human colon cancer
	460	17	0.5	2617	21	AAF18171	Lung cancer associ
c	461	17	0.5	2653	23	ABL19844	Drosophila melanog
	462	17	0.5	2672	16	AAT08693	Cephamycin biosynt
	463	17	0.5	2703	23	ABL11233	Drosophila melanog
	464	17	0.5	2717	19	AAT96741	DNA coding K. oxyt
	465	17	0.5	2723	23	ABL27612	Drosophila melanog
c	466	17	0.5	2745	23	ABL22957	Drosophila melanog
	467	17	0.5	2748	23	ABL09795	Drosophila melanog
c	468	17	0.5	2759	20	AAX28295	Human CYP3A4 gene.
	469	17	0.5	2786	22	AAH34862	Human colon cancer
	470	17	0.5	2786	24	ABL90450	Human polynucleoti
	471	17	0.5	2796	19	AAV22648	Homo sapiens trans
	472	17	0.5	2812	22	AAH17702	Human cDNA sequenc
c	473	17	0.5	2813	15	AAQ73464	Human Egr-2. Homo
c	474	17	0.5	2817	10	AAN91816	EGR2 clone. Homo
	475	17	0.5	2829	22	AAK51576	Human polynucleoti
	476	17	0.5	2833	22	AAF23416	Human SEC7 DNA seq
	477	17	0.5	2835	22	AAK52560	Human polynucleoti
c	478	17	0.5	2849	24	ABN95612	Gene #2110 used to
c	479	17	0.5	2887	24	AAK99403	DNA of APP related
c	480	17	0.5	2910	24	ABK84640	Human cDNA differe
c	481	17	0.5	2910	24	ABN97289	Gene #3787 used to
c	482	17	0.5	2910	24	AAS61864	Lung small cell ca
c	483	17	0.5	2927	22	AAK86823	Human immune/haema
	484	17	0.5	2929	24	AAD31710	Rat SNF1/AMPK-Rela
	485	17	0.5	2930	24	ABQ88181	Human osteoblast d
	486	17	0.5	2930	24	ABK70285	Human lung cancer
c	487	17	0.5	2934	24	ABI99306	Mouse ischaemic co
c	488	17	0.5	2980	24	AAS94915	Human DNA sequence
	489	17	0.5	3044	24	ABQ88180	Human osteoblast d
	490	17	0.5	3044	24	ABK84552	Human cDNA differe
	491	17	0.5	3044	24	ABN97223	Gene #3721 used to
	492	17	0.5	3047	24	ABK09792	Human ovarian tumo
	493	17	0.5	3072	24	ABQ88182	Human osteoblast d
	494	17	0.5	3115	21	AAC98113	Human colon cancer
	495	17	0.5	3160	19	AAV46392	Human tumour suppr
	496	17	0.5	3160	20	AAX57671	Human tumour suppr
	497	17	0.5	3160	24	ABK84480	Human cDNA differe
	498	17	0.5	3160	24	ABL59528	Human phosphatase
c	499	17	0.5	3169	22	AAI58822	Human polynucleoti
	500	17	0.5	3184	22	ABA09269	Human IGFALS homol
	501	17	0.5	3203	22	AAI60608	Human polynucleoti
	502	17	0.5	3217	12	AAQ12002	Full-length rat an
c	503	17	0.5	3317	23	ABV24474	Human prostate exp
	504	17	0.5	3330	16	AAQ88687	Human very low den
	505	17	0.5	3349	20	AAX60577	Human disintegrin
	506	17	0.5	3402	24	AAD35330	Human toll like re
	507	17	0.5	3410	22	AAH45399	Human ADAM10 codin
	508	17	0.5	3410	22	AAF83470	Human ADAM10 DNA.
c	509	17	0.5	3460	22	AAH72627	Human cervical can
c	510	17	0.5	3460	22	AAH72735	Human cervical can
c	511	17	0.5	3504	24	ABL34314	Human immune syste
	512	17	0.5	3523	23	ABL09879	Drosophila melanog
c	513	17	0.5	3531	21	AAC58368	Human PRO290 nucle
c	514	17	0.5	3531	21	AAZ64954	Membrane-bound pro
c	515	17	0.5	3531	22	AAF44100	Human PRO290 (UNQ2
c	516	17	0.5	3546	22	AAF72802	Secreted protein g
	517	17	0.5	3597	23	ABL09813	Drosophila melanog
	518	17	0.5	3621	23	ABL17482	Drosophila melanog
	519	17	0.5	3656	17	AAT36751	VLDL receptor gene
	520	17	0.5	3773	23	ABL11845	Drosophila melanog

	521	17	0.5	3782	23	ABL28681	Drosophila melanog
c	522	17	0.5	3969	24	ABN95715	Gene #2213 used to
c	523	17	0.5	3978	23	AAS87429	DNA encoding novel
c	524	17	0.5	4010	16	AAQ84303	Human NF-AT120, X
	525	17	0.5	4138	24	AAI72578	Human Na+-driven C
c	526	17	0.5	4159	23	ABL10310	Drosophila melanog
c	527	17	0.5	4171	22	AAH18676	Human cDNA sequenc
	528	17	0.5	4188	22	AAF98721	Human late stage o
	529	17	0.5	4250	24	ABK86098	Human cDNA for hae
c	530	17	0.5	4311	23	ABL01926	Drosophila melanog
c	531	17	0.5	4321	24	ABI99593	Mouse ischaemic co
	532	17	0.5	4351	18	AAT97850	M. halophilus ecto
	533	17	0.5	4373	23	ABL28398	Drosophila melanog
	534	17	0.5	4389	23	ABL14960	Drosophila melanog
	535	17	0.5	4389	23	ABL20288	Drosophila melanog
	536	17	0.5	4391	23	ABL20290	Drosophila melanog
c	537	17	0.5	4402	24	ABK12867	cDNA encoding huma
	538	17	0.5	4445	22	AAH18522	Human cDNA sequenc
c	539	17	0.5	4534	20	AAX24069	Human 53BP2 DNA.
c	540	17	0.5	4534	21	AAA09324	Human cancer assoc
	541	17	0.5	4641	21	AAF16251	Human prostate can
	542	17	0.5	4751	22	AAS26783	Human genomic DNA
	543	17	0.5	4753	23	ABL15906	Drosophila melanog
c	544	17	0.5	4957	24	ABN95656	Gene #2154 used to
c	545	17	0.5	4957	24	ABL68507	Kidney cancer rela
c	546	17	0.5	4957	24	ABL68869	Kidney cancer rela
	547	17	0.5	4980	23	ABL11232	Drosophila melanog
	548	17	0.5	5045	23	ABL09794	Drosophila melanog
	549	17	0.5	5128	21	AAA75835	DNA encoding the h
c	550	17	0.5	5310	24	ABL70417	Chemically treated
c	551	17	0.5	5310	24	ABK31444	Signal transductio
c	552	17	0.5	5378	24	ABL33684	Human immune syste
	553	17	0.5	5385	24	AAI72577	Murine Na+-driven
c	554	17	0.5	5404	23	ABL14953	Drosophila melanog
c	555	17	0.5	5467	24	ABL34109	Human immune syste
c	556	17	0.5	5499	24	ABQ66972	Human angiogenesis
	557	17	0.5	5631	23	ABL17158	Drosophila melanog
	558	17	0.5	5749	23	ABL22956	Drosophila melanog
c	559	17	0.5	5798	24	ABQ67012	Human angiogenesis
c	560	17	0.5	5798	24	ABL32169	Human immune syste
c	561	17	0.5	5814	24	ABN59737	Novel human coding
	562	17	0.5	5887	24	ABQ71051	Listeria monocytog
c	563	17	0.5	5997	24	ABL70285	Chemically treated
c	564	17	0.5	5997	24	AAS61210	Human gene regulat
c	565	17	0.5	5997	24	ABK31310	Signal transductio
	566	17	0.5	6065	24	ABL70579	Chemically treated
	567	17	0.5	6065	24	AAS61260	Human gene regulat
	568	17	0.5	6065	24	ABK31356	Signal transductio
c	569	17	0.5	6223	23	ABL14690	Drosophila melanog
	570	17	0.5	6307	24	ABL33341	Human immune syste
c	571	17	0.5	6334	22	AAS46458	Tumour suppressor
	572	17	0.5	6575	24	ABL58984	AAV expression vec
	573	17	0.5	6699	24	ABN80183	Human chemically m
c	574	17	0.5	6731	24	ABL33061	Human immune syste
	575	17	0.5	6734	24	ABK92180	Prostate cancer-as
c	576	17	0.5	6954	24	ABL33390	Human immune syste
c	577	17	0.5	6977	22	AAS46627	Tumour suppressor
c	578	17	0.5	7503	24	ABQ67009	Human angiogenesis
c	579	17	0.5	7503	24	ABL33548	Human immune syste
c	580	17	0.5	7511	24	ABL33283	Human immune syste
	581	17	0.5	7625	22	AAL04533	Human reproductive
	582	17	0.5	7625	23	ABL97456	Human testicular a
	583	17	0.5	7633	22	AAL04536	Human reproductive
	584	17	0.5	7633	23	ABL97459	Human testicular a
c	585	17	0.5	7667	22	AAS46334	Tumour suppressor
c	586	17	0.5	7669	22	AAS46597	Tumour suppressor
c	587	17	0.5	7669	24	ABL33846	Human immune syste
c	588	17	0.5	7900	22	AAS46759	Tumour suppressor
	589	17	0.5	8093	21	AAA81512	N. meningitidis pa
c	590	17	0.5	8451	23	ABL11062	Drosophila melanog
	591	17	0.5	8711	22	AAS26809	Human genomic DNA
c	592	17	0.5	8771	24	ABL33825	Human immune syste
	593	17	0.5	8801	22	AAS45436	Chemically pretrea

	594	17	0.5	8801	24	ABL33740	Human immune syste
	595	17	0.5	8801	24	ABK28285	DNA transcription
c	596	17	0.5	8841	22	AAK78951	Human immune/haema
c	597	17	0.5	8841	22	AAK78952	Human immune/haema
c	598	17	0.5	8919	23	ABL20699	Drosophila melanog
c	599	17	0.5	8968	23	ABL13297	Drosophila melanog
c	600	17	0.5	9002	23	ABL09604	Drosophila melanog
c	601	17	0.5	9072	18	AAV74356	Staphylococcus aur
c	602	17	0.5	9163	24	ABL45808	Human endothelial
	603	17	0.5	9592	17	AAT36752	Adenovirus vector
	604	17	0.5	9990	22	AAH26123	Mouse prion protei
	605	17	0.5	9990	22	AAH26136	Mouse PrP/human ta
	606	17	0.5	10250	24	ABN80082	Human chemically m
	607	17	0.5	10336	22	ABA21257	Human nervous syst
c	608	17	0.5	10394	23	ABL11060	Drosophila melanog
c	609	17	0.5	10682	24	ABN80085	Human chemically m
	610	17	0.5	10747	20	AAX60207	Sequence used to p
	611	17	0.5	10747	21	AAA52319	Human keratin K1 (
c	612	17	0.5	11049	24	ABL92218	Chemically treated
c	613	17	0.5	11049	24	ABL49321	Human polynucleoti
c	614	17	0.5	11049	24	ABL32668	Human immune syste
c	615	17	0.5	11500	23	ABL28680	Drosophila melanog
c	616	17	0.5	11888	22	AAL07198	Human reproductive
c	617	17	0.5	12409	24	AAS63312	Chemically pretrea
	618	17	0.5	12550	23	ABL29300	Drosophila melanog
	619	17	0.5	12870	23	ABL14952	Drosophila melanog
	620	17	0.5	13321	22	AAS46421	Tumour suppressor
	621	17	0.5	13332	23	ABL09812	Drosophila melanog
c	622	17	0.5	13960	23	AAS87164	DNA encoding novel
	623	17	0.5	14093	23	ABK42857	Genomic sequence #
c	624	17	0.5	14170	23	ABL11044	Drosophila melanog
	625	17	0.5	14372	23	ABL09878	Drosophila melanog
	626	17	0.5	14708	23	ABL13296	Drosophila melanog
	627	17	0.5	16818	22	ABA15470	Human nervous syst
	628	17	0.5	16818	22	ABA16242	Human nervous syst
	629	17	0.5	16818	22	ABA18914	Human nervous syst
c	630	17	0.5	16842	22	AAS46411	Tumour suppressor
c	631	17	0.5	16842	24	ABL70383	Chemically treated
c	632	17	0.5	16842	24	AAS61335	Human gene regulat
c	633	17	0.5	16842	24	ABK31418	Signal transductio
c	634	17	0.5	17528	24	ABL32600	Human immune syste
c	635	17	0.5	17893	24	ABL33364	Human immune syste
c	636	17	0.5	17951	23	ABL11844	Drosophila melanog
c	637	17	0.5	18683	24	ABL54333	Chemically treated
c	638	17	0.5	18683	24	ABL32312	Human immune syste
c	639	17	0.5	18871	23	ABK42430	Genomic sequence #
c	640	17	0.5	19233	24	ABL49345	Human polynucleoti
	641	17	0.5	20486	23	ABL20698	Drosophila melanog
	642	17	0.5	22275	23	ABL25566	Drosophila melanog
	643	17	0.5	24978	20	AAX60209	SEQ ID 3 of US5914
	644	17	0.5	24979	21	AAA52321	Genomic DNA sequen
	645	17	0.5	26657	24	AAI72317	Human transporter
c	646	17	0.5	26997	22	AAS46748	Tumour suppressor
	647	17	0.5	28730	21	AAA81486	N. meningitidis pa
c	648	17	0.5	32134	22	AAI99172	Human excretory re
c	649	17	0.5	32134	22	AAI63522	Human kidney relat
c	650	17	0.5	32192	22	AAI99173	Human excretory re
c	651	17	0.5	32192	22	AAI63523	Human kidney relat
c	652	17	0.5	35100	22	AAK65700	Human immune/haema
c	653	17	0.5	35100	22	AAK69767	Human immune/haema
c	654	17	0.5	35115	22	AAK65699	Human immune/haema
c	655	17	0.5	35115	22	AAK69766	Human immune/haema
c	656	17	0.5	35465	22	AAF54723	Nucleotide sequenc
	657	17	0.5	37856	21	AAA11992	S. cellulosum DNA
c	658	17	0.5	37973	24	ABL34197	Human immune syste
c	659	17	0.5	48037	22	AAK84729	Human immune/haema
c	660	17	0.5	48037	22	AAK85983	Human immune/haema
c	661	17	0.5	48045	22	AAK84730	Human immune/haema
c	662	17	0.5	48045	22	AAK85984	Human immune/haema
c	663	17	0.5	56485	21	AAA81476	N. meningitidis pa
	664	17	0.5	63563	22	AAF28546	Genomic fragment #
c	665	17	0.5	72750	21	AAA81468	N. meningitidis pa
	666	17	0.5	73947	23	ABL07230	Drosophila melanog

c	667	17	0.5	83391	24	ABQ67094	Human angiogenesis
	668	17	0.5	83946	24	ABQ88101	Human osteoblast d
	669	17	0.5	349980	21	AAF21544	Neisseria meningit
c	670	17	0.5	349980	21	AAF21544	Neisseria meningit
c	671	17	0.5	349980	21	AAF21609	Neisseria meningit
c	672	17	0.5	349980	21	AAF21610	Neisseria meningit
c	673	17	0.5	349980	22	AAH41223	Pyrococcus abyssi
c	674	17	0.5	349980	22	AAH41224	Pyrococcus abyssi
c	675	17	0.5	349980	22	AAH41225	Pyrococcus abyssi
c	676	17	0.5	349980	22	AAH68533	C glutamicum codin
	677	17	0.5	368004	24	ABL57909	Human transporter
	678	17	0.5	837096	21	AAA81489	N. meningitidis pa
c	679	17	0.5	1437668	21	AAA81490	N. meningitidis B
	680	17	0.5	1664976	19	AAV21209	Methanococcus jann
	681	17	0.5	2365589	24	ABA90521	Genomic sequence o
	682	17	0.5	4403765	22	AAI99683	Mycobacterium tube
c	683	17	0.5	4403765	22	AAI99683	Mycobacterium tube
	684	17	0.5	4411529	22	AAI99682	Mycobacterium tube
c	685	17	0.5	4411529	22	AAI99682	Mycobacterium tube
c	686	16	0.5	18	20	AAZ31863	Human G-alpha-13 a
	687	16	0.5	20	22	AAF91298	Human E2F transcri
	688	16	0.5	25	19	AAV07946	Helicobacter pylor
	689	16	0.5	26	19	AAV07943	Helicobacter pylor
c	690	16	0.5	26	21	AAZ55457	Neisseria species
c	691	16	0.5	26	22	AAS44219	Neisseria meningit
c	692	16	0.5	26	22	AAS44241	Neisseria meningit
c	693	16	0.5	26	22	AAS44243	Neisseria meningit
c	694	16	0.5	26	22	AAS44245	Neisseria meningit
c	695	16	0.5	26	22	AAS44247	Neisseria meningit
c	696	16	0.5	26	22	AAS44344	Neisseria meningit
c	697	16	0.5	26	22	AAD17062	N. meningitidis fu
c	698	16	0.5	26	22	AAD17107	N. meningitidis (d
	699	16	0.5	31	22	AAS44240	Neisseria meningit
	700	16	0.5	31	22	AAS44244	Neisseria meningit
c	701	16	0.5	35	22	AAS44249	Neisseria meningit
	702	16	0.5	47	21	AAZ68103	Human map-related
c	703	16	0.5	47	24	AAL47108	Peptide presentati
c	704	16	0.5	47	24	AAL47109	Peptide presentati
c	705	16	0.5	50	21	AAC26760	Human secreted pro
	706	16	0.5	50	22	AAD15026	Oryctolagus cunicu
	707	16	0.5	51	22	AAL28808	Human SNP oligonuc
	708	16	0.5	54	21	AAA73936	GFP Gly(GGC)5 forw
c	709	16	0.5	54	21	AAA73955	GFP Pro(CCG)5 forw
	710	16	0.5	57	20	AAX60226	Fragment of the PA
c	711	16	0.5	60	24	ABN34980	Human spliced tran
c	712	16	0.5	60	24	ABN42530	Human spliced tran
c	713	16	0.5	65	24	ABN56503	Mouse spliced tran
c	714	16	0.5	67	16	AAQ79642	Nucleotide sequenc
	715	16	0.5	99	22	ABA47799	Human breast cell
	716	16	0.5	99	22	AAI06347	Probe #6338 used t
	717	16	0.5	99	24	ABS13932	Human genome-deriv
c	718	16	0.5	105	21	AAC18533	Human secreted pro
	719	16	0.5	119	19	AAV16155	Microsatellite mar
	720	16	0.5	126	21	AAA31060	Plant microsatelli
	721	16	0.5	145	17	AAT28159	Senescence-related
c	722	16	0.5	146	24	ABS17316	Human genome-deriv
	723	16	0.5	152	19	AAV16158	Microsatellite mar
c	724	16	0.5	152	21	AAH51644	Sbg1 exon sequence
c	725	16	0.5	152	21	AAH51667	Sbg1 exon sequence
c	726	16	0.5	152	21	AAH51679	Sbg1 exon sequence
c	727	16	0.5	152	21	AAC04711	Human secreted pro
	728	16	0.5	159	22	AAF84452	Chicken beta-actin
	729	16	0.5	160	22	AAK60781	Human immune/haema
	730	16	0.5	166	20	AAX01048	Septoria sequence
c	731	16	0.5	171	21	AAA31044	Plant microsatelli
	732	16	0.5	179	22	AAL19730	Human breast cance
	733	16	0.5	186	21	AAC13994	Human secreted pro
	734	16	0.5	192	21	AAA81855	N. meningitidis pa
	735	16	0.5	195	21	AAA31565	Plant microsatelli
	736	16	0.5	200	21	AAC25887	Human secreted pro
c	737	16	0.5	204	23	ABL22711	Drosophila melanog
	738	16	0.5	208	21	AAC12643	Human secreted pro
	739	16	0.5	210	22	ABA69986	Human foetal liver

	740	16	0.5	210	22	ABA36815	Probe #15281 for g
	741	16	0.5	210	22	AAK18195	Human brain expres
	742	16	0.5	210	22	AAK44089	Human bone marrow
	743	16	0.5	210	22	AAI24723	Probe #14656 for g
	744	16	0.5	210	22	AAI50098	Probe #18784 used
	745	16	0.5	210	24	ABS18323	Human genome-deriv
c	746	16	0.5	216	24	ABN25051	Human ORFX polynuc
	747	16	0.5	217	24	ABN25563	Human ORFX polynuc
	748	16	0.5	225	21	AAC19040	Human secreted pro
c	749	16	0.5	229	16	AAT22982	Human gene signatu
	750	16	0.5	229	21	AAC28342	Human secreted pro
c	751	16	0.5	230	22	ABA73231	Human foetal liver
c	752	16	0.5	230	22	ABA38655	Probe #17121 for g
c	753	16	0.5	230	22	AAK21666	Human brain expres
c	754	16	0.5	230	22	AAK47831	Human bone marrow
c	755	16	0.5	230	22	AAI26055	Probe #15988 for g
c	756	16	0.5	230	22	AAI53659	Probe #22345 used
c	757	16	0.5	230	24	ABS21830	Human genome-deriv
	758	16	0.5	235	21	AAC30438	Human secreted pro
c	759	16	0.5	243	22	AAS29862	Human lung antigen
c	760	16	0.5	250	21	AAC18381	Human secreted pro
c	761	16	0.5	256	20	AAV88334	EST clone FX194.
	762	16	0.5	258	21	AAC68906	Rat ADIPO7 coding
c	763	16	0.5	263	21	AAA31075	Plant microsatelli
c	764	16	0.5	267	22	AAS28361	Genomic sequence #
	765	16	0.5	271	21	AAC06348	Human secreted pro
c	766	16	0.5	271	22	AAS27956	Novel cDNA encodin
	767	16	0.5	274	24	ABL71358	Corn tassel-derive
	768	16	0.5	276	24	ABL71230	Corn tassel-derive
	769	16	0.5	283	24	ABK62781	Rat sequence diffe
c	770	16	0.5	284	13	AAQ22701	Sequence encoding
c	771	16	0.5	284	24	ABN95468	Gene #1966 used to
c	772	16	0.5	284	24	ABK64465	Human benign prost
c	773	16	0.5	284	24	ABL68159	Kidney cancer rela
	774	16	0.5	286	24	ABN21304	Human ORFX polynuc
	775	16	0.5	287	22	AAL11882	Human breast cance
	776	16	0.5	287	22	AAL20772	Human breast cance
c	777	16	0.5	287	22	AAK54240	Murine translation
c	778	16	0.5	287	24	ABL37093	Human colon tumour
	779	16	0.5	289	22	AAL20254	Human breast cance
c	780	16	0.5	289	24	ABN62942	Human cancer relat
	781	16	0.5	290	24	ABK31117	Plant dwarfing/stu
c	782	16	0.5	292	24	ABN77115	Human ORF2062 cDNA
	783	16	0.5	298	20	AAX40983	Human secreted pro
c	784	16	0.5	300	20	AAZ14673	Human gene express
	785	16	0.5	300	21	AAA01470	Human colon cancer
c	786	16	0.5	301	21	AAC06635	Human secreted pro
	787	16	0.5	301	24	ABK79104	Bacillus clausii g
c	788	16	0.5	304	24	ABL71158	Corn tassel-derive
c	789	16	0.5	307	21	AAC40720	Zea mays DNA fragm
	790	16	0.5	310	24	ABL81208	Human ovarian canc
	791	16	0.5	313	22	AAS39057	Novel human diagno
c	792	16	0.5	314	20	AAZ31890	Mouse Dlx5 5'untra
	793	16	0.5	314	20	AAV89554	EST clone CR1190.
c	794	16	0.5	314	24	ABS20910	Human genome-deriv
c	795	16	0.5	315	21	AAC98270	Human colon cancer
c	796	16	0.5	318	21	AAA31182	Plant microsatelli
	797	16	0.5	321	22	AAD13973	Human 5' UTR of FM
c	798	16	0.5	322	18	AAT75508	P. americanus anti
c	799	16	0.5	322	21	AAZ49961	Winter flounder li
	800	16	0.5	323	21	AAC27096	Human secreted pro
c	801	16	0.5	330	24	ABN20027	Human ORFX polynuc
	802	16	0.5	339	22	AAI83366	Human polynucleoti
c	803	16	0.5	339	24	AAL43490	Human stresscopin
c	804	16	0.5	342	14	AAQ61118	Human brain Expres
c	805	16	0.5	345	24	ABL88684	Conus striatus I-s
	806	16	0.5	348	22	AAS07101	DNA encoding Group
	807	16	0.5	350	21	AAC56190	Eucalyptus grandis
	808	16	0.5	351	23	AAS57713	cDNA #389 encoding
	809	16	0.5	352	23	ABV56593	Human prostate exp
	810	16	0.5	355	22	AAI29472	Colon tumour relat
c	811	16	0.5	359	21	AAC56476	Eucalyptus grandis
	812	16	0.5	362	24	ABN18066	Human ORFX polynuc

c	813	16	0.5	365	16	AAT19524	Human gene signatu
c	814	16	0.5	367	22	ABA47063	Human breast cell
c	815	16	0.5	367	22	ABA64946	Human foetal liver
c	816	16	0.5	367	22	ABA32055	Probe #10521 for g
c	817	16	0.5	367	22	AAK13373	Human brain expres
c	818	16	0.5	367	22	AAK39110	Human bone marrow
c	819	16	0.5	367	22	AAI19919	Probe #9852 for ge
c	820	16	0.5	367	22	AAI45114	Probe #13800 used
c	821	16	0.5	367	22	AAI05633	Probe #5624 used t
c	822	16	0.5	367	24	ABS13192	Human genome-deriv
c	823	16	0.5	367	24	ABN20251	Human ORFX polynuc
c	824	16	0.5	369	24	ABN76446	Human ORF1393 cDNA
	825	16	0.5	372	20	AAV73008	Human adult placen
	826	16	0.5	372	23	AAS68870	DNA encoding novel
	827	16	0.5	372	24	ABQ92033	Human polynucleoti
	828	16	0.5	373	21	AAC22865	Human secreted pro
	829	16	0.5	373	22	ABA58544	Human foetal liver
	830	16	0.5	373	22	AAK06663	Human brain expres
	831	16	0.5	373	22	AAK32359	Human bone marrow
	832	16	0.5	373	22	AAI38206	Probe #6892 used t
	833	16	0.5	373	24	ABS07142	Human genome-deriv
c	834	16	0.5	378	24	ABN24549	Human ORFX polynuc
c	835	16	0.5	378	24	ABL81730	Human ovarian canc
c	836	16	0.5	379	21	AAA31636	Plant microsatelli
c	837	16	0.5	381	23	ABL41662	Nucleotide sequenc
	838	16	0.5	381	24	ABN21464	Human ORFX polynuc
	839	16	0.5	384	20	AAX40311	Human secreted pro
	840	16	0.5	384	22	AAI81797	Human polynucleoti
	841	16	0.5	385	21	AAC02615	Human secreted pro
	842	16	0.5	385	23	AAS69852	DNA encoding novel
	843	16	0.5	386	18	AAT84649	RFHV glycoprotein
	844	16	0.5	387	22	AAI29173	Colon tumour relat
	845	16	0.5	387	22	AAF66682	Novel human polynu
c	846	16	0.5	387	22	AAF67407	Novel human polynu
	847	16	0.5	390	15	AAQ73668	Tick allergen Derf
	848	16	0.5	390	15	AAQ73671	Modified major tic
	849	16	0.5	390	15	AAQ73672	Modified major tic
	850	16	0.5	390	15	AAQ73670	Tick allergen Derf
	851	16	0.5	390	16	AAQ89988	Modified Derf II m
	852	16	0.5	390	16	AAQ89989	Modified Derf II m
	853	16	0.5	390	16	AAQ89990	Modified Derf II m
	854	16	0.5	390	16	AAQ89991	Modified Derf II m
	855	16	0.5	390	16	AAQ89992	Modified Derf II m
	856	16	0.5	390	16	AAQ89993	Modified Derf II m
	857	16	0.5	390	16	AAQ89994	Modified Derf II m
	858	16	0.5	390	16	AAQ89983	Modified Derf II m
	859	16	0.5	390	16	AAQ89984	Modified Derf II m
	860	16	0.5	390	16	AAQ89985	Modified Derf II m
	861	16	0.5	390	16	AAQ89986	Modified Derf II m
	862	16	0.5	390	16	AAQ89987	Modified Derf II m
	863	16	0.5	390	16	AAQ89978	Modified Derf II m
	864	16	0.5	390	16	AAQ89979	Modified Derf II m
	865	16	0.5	390	16	AAQ89980	Modified Derf II m
	866	16	0.5	390	16	AAQ89981	Modified Derf II m
	867	16	0.5	390	16	AAQ89967	Modified Derf II m
	868	16	0.5	390	16	AAQ89968	Modified Derf II m
	869	16	0.5	390	16	AAQ89969	Modified Derf II m
	870	16	0.5	390	16	AAQ89970	Modified Derf II m
	871	16	0.5	390	16	AAQ89971	Modified Derf II m
	872	16	0.5	390	16	AAQ89972	Modified Derf II m
	873	16	0.5	390	16	AAQ89973	Modified Derf II m
	874	16	0.5	390	16	AAQ89974	Modified Derf II m
	875	16	0.5	390	16	AAQ89975	Modified Derf II m
	876	16	0.5	390	16	AAQ89976	Modified Derf II m
	877	16	0.5	390	16	AAQ89982	Modified Derf II m
	878	16	0.5	390	16	AAQ89977	Modified Derf II m
	879	16	0.5	390	16	AAQ89966	Modified Derf II m
	880	16	0.5	390	16	AAQ89963	Native Derf II maj
	881	16	0.5	390	16	AAQ89965	Native Derf II maj
	882	16	0.5	390	17	AAT41953	Modified major tic
	883	16	0.5	390	17	AAT41954	Modified major tic
	884	16	0.5	390	17	AAT41955	Modified major tic
	885	16	0.5	390	17	AAT41956	Modified major tic
	886	16	0.5	390	17	AAT41957	Modified major tic
	887	16	0.5	390	17	AAT41949	Major tick allerge
	888	16	0.5	390	17	AAT41950	Major tick allerge
	889	16	0.5	390	17	AAT41951	Modified major tic
	890	16	0.5	390	17	AAT41952	Modified major tic
	891	16	0.5	390	17	AAT41958	Modified major tic
	892	16	0.5	390	17	AAT41959	Modified major tic
	893	16	0.5	390	17	AAT41960	Modified major tic
	894	16	0.5	390	22	AAI89540	Human polynucleoti
	895	16	0.5	393	15	AAQ54064	Der f II Tick majo
	896	16	0.5	393	22	AAF67438	Novel human polynu
	897	16	0.5	396	22	AAI81388	Human polynucleoti
c	898	16	0.5	396	22	AAF94816	Human ovarian canc
c	899	16	0.5	396	22	AAF94872	Human ovarian canc
c	900	16	0.5	396	22	AAF94873	Human ovarian canc
c	901	16	0.5	396	24	ABT03083	Human ovarian carc
c	902	16	0.5	396	24	ABT03139	Human ovarian carc
c	903	16	0.5	396	24	ABT03140	Human ovarian carc
c	904	16	0.5	396	24	ABL48766	Ovarian carcinoma
c	905	16	0.5	396	24	ABL48822	Ovarian carcinoma
c	906	16	0.5	396	24	ABL48823	Ovarian carcinoma
c	907	16	0.5	399	24	ABQ54857	Human ovarian anti
	908	16	0.5	399	24	ABN73862	Bovine embryonic g
c	909	16	0.5	399	24	AAD33401	Human urocortin-re
	910	16	0.5	400	18	AAV78347	Staphylococcus aur
	911	16	0.5	401	22	AAK96048	Human neuregulin g
	912	16	0.5	401	22	AAK96049	Human neuregulin g
	913	16	0.5	401	22	AAK97541	Human neuregulin g
	914	16	0.5	401	22	AAK97542	Human neuregulin g
	915	16	0.5	401	22	AAH99146	Human EST-derived
	916	16	0.5	402	23	AAD23448	Human lung tumour-
	917	16	0.5	403	21	AAC09406	Human secreted pro
	918	16	0.5	404	22	AAH42798	Nucleotide sequenc
c	919	16	0.5	406	22	ABA45369	Human breast cell
c	920	16	0.5	406	22	ABA55856	Human foetal liver
c	921	16	0.5	406	22	ABA25531	Probe #3997 for ge
c	922	16	0.5	406	22	AAK04075	Human brain expres
c	923	16	0.5	406	22	AAK29558	Human bone marrow
c	924	16	0.5	406	22	AAI14130	Probe #4063 for ge
c	925	16	0.5	406	22	AAI35513	Probe #4199 used t
c	926	16	0.5	406	22	AAI03980	Probe #3971 used t
c	927	16	0.5	406	24	ABS04112	Human genome-deriv
c	928	16	0.5	409	22	AAI80936	Human polynucleoti
c	929	16	0.5	411	22	AAF66875	Novel human polynu
	930	16	0.5	412	22	ABA58817	Human foetal liver
	931	16	0.5	412	22	AAK06971	Human brain expres
	932	16	0.5	412	22	AAK32707	Human bone marrow
	933	16	0.5	412	22	AAI38522	Probe #7208 used t
c	934	16	0.5	412	23	ABV61956	Human prostate exp
	935	16	0.5	412	24	ABS07503	Human genome-deriv
	936	16	0.5	416	21	AAC79207	Human lung tumour-
	937	16	0.5	416	23	AAD23283	Human lung tumour-
	938	16	0.5	418	22	AAL11355	Human breast cance
c	939	16	0.5	419	21	AAA31136	Plant microsatelli
c	940	16	0.5	420	21	AAC41346	Zea mays DNA fragm
	941	16	0.5	423	22	AAL10150	Human breast cance
	942	16	0.5	424	22	AAI81908	Human polynucleoti
c	943	16	0.5	424	23	AAS57348	cDNA #24 encoding
	944	16	0.5	424	24	ABL87144	Human ovarian canc
	945	16	0.5	425	21	AAC43732	Zea mays DNA fragm
	946	16	0.5	427	21	AAZ51838	Maize plastid targ
	947	16	0.5	428	24	ABK53867	Human head and nec
	948	16	0.5	430	24	ABK76055	Bacillus lichenifo
c	949	16	0.5	435	21	AAC56506	Eucalyptus grandis
c	950	16	0.5	439	22	AAL06748	Human reproductive
c	951	16	0.5	439	22	AAI62654	Human breast or ov
c	952	16	0.5	440	21	AAA35334	Myrtaceae microsat
c	953	16	0.5	440	22	AAH06282	Human cDNA clone (
	954	16	0.5	441	21	AAC79262	Human lung tumour-
	955	16	0.5	441	23	AAD23338	Human lung tumour-
	956	16	0.5	445	22	AAK92100	Human cDNA 5'-end
	957	16	0.5	445	22	AAK93630	Human cDNA clone r
	958	16	0.5	447	22	ABA59050	Human foetal liver

	959	16	0.5	447	22	ABA27862	Probe #6328 for ge
	960	16	0.5	447	22	AAK07223	Human brain expres
	961	16	0.5	447	22	AAK32979	Human bone marrow
	962	16	0.5	447	22	AAI16281	Probe #6214 for ge
	963	16	0.5	447	22	AAI38781	Probe #7467 used t
c	964	16	0.5	447	24	ABS04802	Human genome-deriv
	965	16	0.5	447	24	ABS07787	Human genome-deriv
c	966	16	0.5	448	22	ABA20103	Human nervous syst
c	967	16	0.5	451	22	ABA13679	Human nervous syst
c	968	16	0.5	451	22	AAF87079	Rice isoleucyl-RNA
	969	16	0.5	451	24	ABL79824	Human ovarian canc
	970	16	0.5	452	24	ABN75862	Human synthase-lik
c	971	16	0.5	456	21	AAC01184	Human secreted pro
	972	16	0.5	456	22	ABA42594	Human breast cell
	973	16	0.5	456	22	ABA53024	Human foetal liver
	974	16	0.5	456	22	ABA22798	Probe #1264 for ge
	975	16	0.5	456	22	AAK01272	Human brain expres
	976	16	0.5	456	22	AAK26733	Human bone marrow
	977	16	0.5	456	22	AAI11359	Probe #1292 for ge
	978	16	0.5	456	22	AAI32628	Probe #1314 used t
	979	16	0.5	456	22	AAI01275	Probe #1266 used t
	980	16	0.5	456	24	ABS01327	Human genome-deriv
	981	16	0.5	457	16	AAQ95181	Simple tandem repe
	982	16	0.5	457	21	AAC01875	Human secreted pro
c	983	16	0.5	458	22	AAK68444	Human immune/haema
c	984	16	0.5	458	22	AAK84980	Human immune/haema
c	985	16	0.5	461	21	AAC16175	Human secreted pro
c	986	16	0.5	462	24	ABK79828	Bacillus clausii g
c	987	16	0.5	464	21	AAA31842	Plant microsatelli
	988	16	0.5	464	22	ABA57359	Human foetal liver
	989	16	0.5	464	22	ABA26895	Probe #5361 for ge
	990	16	0.5	464	22	AAK05395	Human brain expres
	991	16	0.5	464	22	AAK30992	Human bone marrow
	992	16	0.5	464	22	AAI15496	Probe #5429 for ge
	993	16	0.5	464	22	AAI36903	Probe #5589 used t
	994	16	0.5	464	22	AAH06056	Human cDNA clone (
	995	16	0.5	464	24	ABS05741	Human genome-deriv
c	996	16	0.5	465	24	ABL01416	Murine apoptosis r
	997	16	0.5	468	21	AAC38306	Zea mays DNA fragm
c	998	16	0.5	469	22	AAI33088	Probe #1774 used t
c	999	16	0.5	469	24	ABS08498	Human genome-deriv
c	1000	16	0.5	471	21	AAC44472	Zea mays DNA fragm

ALIGNMENTS

```
RESULT 1
AAA15299
ID   AAA15299 standard; DNA; 3204 BP.
XX
AC   AAA15299;
XX
DT   04-SEP-2000  (first entry)
XX
DE   DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX
KW   Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3204
FT                   /*tag=  a
XX
PN   WO200026375-A2.
XX
PD   11-MAY-2000.
XX
PF   28-OCT-1999;   99WO-FR02643.
XX
PR   30-OCT-1998;   98FR-0013693.
XX
PA   (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Aujame L,  Bouchardon A,  Renauld-Mongenie G,  Rokbi B,  Nassif X;
PI   Tinsley C,  Perrin A;
XX
DR   WPI; 2000-365622/31.
DR   P-PSDB; AAY93268.
XX
PT   New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT   or preventative vaccines and for diagnosis  -
XX
PS   Claim 2; Page 61-65; 187pp; French.
XX
CC   The present sequence encodes a protein that is specific for pathogenic
CC   strains of Neisseria. The polynucleotides, polypeptides, or their
CC   antigenic fragments, are used in vaccines to treat or protect against
CC   Neisseria infections, particularly by N. meningitidis. The
CC   polynucleotide sequence is also used for recombinant production of
CC   the polypeptide and to produce attenuated Neisseria strains that
CC   overexpress it, or express it in a non-toxic mutant form.
XX
SQ   Sequence 3204 BP; 831 A; 909 C; 899 G; 565 T; 0 other;

  Query Match             100.0%;  Score 3204;  DB 21;  Length 3204;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120

Qy        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180

Qy        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240

Qy        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300

Qy        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360

Qy        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420

Qy        421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480

Qy        481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540

Qy        541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600

Qy        601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660

Qy        661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720

Qy     721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780

Qy     781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840

Qy     841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900

Qy     901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960

Qy     961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTCGGCAAGCAATGAC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTCGGCAAGCAATGAC 1020

Qy    1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080

Qy    1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140

Qy    1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200

Qy    1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260

Qy    1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320

Qy    1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380

Qy    1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440

Qy    1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500

Qy    1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560

Qy    1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620

Qy    1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680

Qy    1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740

Qy    1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800

Qy    1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860

Qy    1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920

Qy    1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980

Qy    1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040

Qy    2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100

Qy    2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160

Qy    2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220

Qy    2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280

Qy    2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340

Qy    2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400

Qy    2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460

Qy    2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520

Qy    2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580

Qy    2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640

Qy    2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700

Qy    2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760

Qy    2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820

Qy    2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
```

```
Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180

Qy   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||
Db   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204


RESULT 2
AAZ53301
ID   AAZ53301 standard; DNA; 1365 BP.
XX
AC   AAZ53301;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 140 partial DNA sequence SEQ ID NO:551.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74539.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 396-397; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1365 BP; 319 A; 410 C; 418 G; 218 T; 0 other;

  Query Match             30.9%;  Score 990;  DB 21;  Length 1365;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1340;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   1858 GGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATC 1917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 GGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATC 78

Qy   1918 GGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTC 1977
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     79 GGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGTCTGCTGGCTTCCCTC 138

Qy   1978 GACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGC 2037
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    139 GACAGCGTCGAAAAAACAGCGGGTAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGC 198

Qy   2038 GGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACAC 2097
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 GGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACAC 258

Qy   2098 GCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAA 2157
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 GCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAA 318

Qy   2158 TCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGC 2217
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 TCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGC 378

Qy   2218 ATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCC 2277
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 ATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCC 438

Qy   2278 GACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    439 GACGGTGTACGCATCTTCAACAATCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 498

Qy   2338 CATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCT 2397
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    499 CATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCT 558

Qy   2398 ACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGT 2457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    559 ACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGT 618

Qy   2458 GTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAA 2517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    619 GTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAA 678

Qy   2518 AATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCA 2577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    679 AATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCA 738

Qy   2578 AATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATC 2637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    739 AATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATC 798

Qy   2638 GGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACC 2697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      799 GGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACC 858

Qy     2698 GGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTG 2757
            ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db      859 GGTGCGGACGAACATGCGGAAGCGAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTG 918

Qy     2758 GGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGC 2817
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      919 GGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGC 978

Qy     2818 TACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGC 2877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      979 TACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGC 1038

Qy     2878 AACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTG 2937
            |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1039 AACAGCATCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTG 1098

Qy     2938 AGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGAC 2997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1099 AGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGAC 1158

Qy     2998 TACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGC 3057
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1159 TACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGC 1218

Qy     3058 AATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGC 3117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1219 AATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGC 1278

Qy     3118 TGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA 3177
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1279 TGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA 1338

Qy     3178 CGAGTCGGCGTAGGCTACCGGTTCTGA 3204
            |||||||||||||||||||||||||||
Db     1339 CGAGTCGGCGTAGGCTACCGGTTCTGA 1365
```

```
RESULT 3
AAZ53300
ID   AAZ53300 standard; DNA; 1365 BP.
XX
AC   AAZ53300;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 140 partial DNA sequence SEQ ID NO:549.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,   Galeotti C,   Grandi G,   Hickey E,   Masignani V,   Mora M;
PI   Petersen J,   Pizza M,   Rappuoli R,   Ratti G,   Scalato E,   Scarselli M;
PI   Tettelin H,   Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74538.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 395; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1365 BP; 317 A; 413 C; 421 G; 214 T; 0 other;

  Query Match             26.0%;  Score 834;  DB 21;  Length 1365;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 183

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      184 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 243

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 303

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      304 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 363

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 423

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      424 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 483

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db      484 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 543

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db      544 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 603

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      604 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 663

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      664 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 723

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 783

Qy       2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 843

Qy       2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 903

Qy       2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 963

Qy       2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 1023

Qy       2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1024 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 1083

Qy       2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 1143

Qy       2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 1203

.Qy      3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
              |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db       1204 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 1263

Qy       3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 1323

Qy       3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
              ||||||||||||||||||||||||||||||||||||||||||
Db       1324 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 1365


RESULT 4
AAS43905
ID   AAS43905 standard; DNA; 4218 BP.
XX
AC   AAS43905;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein 961cL-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27608.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 23; Page 75-76; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;

  Query Match             26.0%;  Score 834;  DB 22;  Length 4218;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2971 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3030

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3031 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3090

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3091 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3150

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       3151 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3210

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3211 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3270

Qy       2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3271 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3330

Qy       2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       3331 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3390

Qy       2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db       3391 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3450

Qy       2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3451 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3510

Qy       2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3511 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3570

Qy       2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3571 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3630

Qy       2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3631 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3690
```

```
Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3691 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3750

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3751 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3810

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3811 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3870

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   3871 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3930

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3931 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3990

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3991 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4050

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   4051 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4110

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4111 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4170

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||||||||||||||||||||
Db   4171 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212
```

```
RESULT 5
AAD17058
ID   AAD17058 standard; DNA; 4218 BP.
XX
AC   AAD17058;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 961cL-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   961cL-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4212
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 961cL-983
FT                   fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10043.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 23; Page 33-34; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) 961cL-983
CC   fusion protein.
XX
SQ   Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;

  Query Match             26.0%;  Score 834;  DB 22;  Length 4218;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2971 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3030

Qy   2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3031 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3090

Qy   2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3091 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3150

Qy   2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   3151 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3210

Qy   2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3211 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3270

Qy   2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3271 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3330

Qy   2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   3331 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3390

Qy   2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
          |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   3391 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3450

Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3451 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3510

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3511 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3570

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3571 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3630

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3631 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3690
```

```
Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3691 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3750

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3751 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3810

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3811 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3870

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     3871 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3930

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3931 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3990

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3991 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4050

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     4051 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4110

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4111 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4170

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
            ||||||||||||||||||||||||||||||||||||||||||
Db     4171 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212


RESULT 6
AAA81527/c
ID   AAA81527 standard; DNA; 13423 BP.
XX
AC   AAA81527;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_74 SEQ ID NO:74.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1439-1443; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 13423 BP; 2916 A; 3336 C; 3842 G; 3329 T; 0 other;

  Query Match             26.0%;  Score 834;  DB 21;  Length 13423;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8882 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 8823

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8822 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 8763

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8762 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 8703

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     8702 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 8643

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8642 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 8583

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8582 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 8523

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     8522 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 8463

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     8462 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 8403

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8402 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 8343

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
```

```
Db   8342 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 8283

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8282 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 8223

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8222 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 8163

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8162 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 8103

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8102 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 8043

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8042 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 7983

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   7982 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 7923

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7922 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 7863

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7862 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 7803

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   7802 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 7743

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7742 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 7683

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||||||||||||||||||||
Db   7682 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 7641
```

```
RESULT 7
AAF21612
ID   AAF21612 standard; DNA; 349980 BP.
XX
AC   AAF21612;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,  Hickey E,  Peterson J,  Tettelin H,  Venter JC,  Masignani V;
PI   Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V,  Rappuoli R;
PI   Frazer CM,  Grandi G;
XX
DR   WPI; 2000-647603/62.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own it was split into 8
CC   sequences which overlap each other at the beginning and end of each
CC   sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC   the beginning  of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC   the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC   Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC   AAF21606 represent PCR primers which are used in the exemplification of
CC   the present invention. The NMB genome and fragments from it have
CC   antibacterial activity, and can be used in vaccines and gene therapy.
CC   Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC   proteins can be used in compositions for treating or preventing infection
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

  Query Match             26.0%;  Score 834;  DB 21;  Length 349980;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1234;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263419 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 263478

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263479 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 263538

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263539 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 263598

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   263599 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 263658

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263659 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 263718

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263719 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 263778

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   263779 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 263838

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 263839 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 263898

Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263899 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 263958

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db 263959 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 264018

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264019 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 264078

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264079 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 264138

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264139 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 264198

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264199 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 264258

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264259 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 264318

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db 264319 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 264378

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264379 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 264438

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264439 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 264498

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db 264499 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 264558

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264559 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 264618

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||||||||||||||||||||
Db 264619 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 264660
```

```
RESULT 8
AAS43875
ID   AAS43875 standard; DNA; 3939 BP.
XX
AC   AAS43875;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-741 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27576.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 44-45; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 3939 BP; 1046 A; 1100 C; 1107 G; 686 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy       2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy       2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy       2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy       2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy       2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy       2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy       2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db       2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy       2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445
```

```
Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||||||||||||||||||||
Db    3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

RESULT 9
AAS43880

```
ID   AAS43880 standard; DNA; 3939 BP.
XX
AC   AAS43880;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G741-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27581.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 51-52; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 3939 BP; 1046 A; 1103 C; 1106 G; 684 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2674 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2733

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2734 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2793

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2794 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2853

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    2854 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2913

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2914 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2973

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2974 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3033

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    3034 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3093

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           ||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db    3094 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3153

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3154 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3213

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    3214 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3273

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    3274 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3333

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3334 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3393

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3394 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3453

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3454 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3513

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3514 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3573

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    3574 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3633

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3634 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3693

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3694 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3753

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    3754 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3813

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3814 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3873

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||||||||||||||||||||
Db    3874 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912
```

RESULT 10

```
AAD17040
ID   AAD17040 standard; DNA; 3939 BP.
XX
AC   AAD17040;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-741 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G983-741 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3939
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G983-741 fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10023.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 3; Page 15-16; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-741
CC   fusion protein.
XX
SQ   Sequence 3939 BP; 1046 A; 1100 C; 1107 G; 686 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565
```

```
Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||
Db     2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 11
AAD17045
ID   AAD17045 standard; DNA; 3939 BP.
XX
AC   AAD17045;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G741-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G741-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3939
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G741-983 fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10030.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 4; Page 20-21; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G741-983
CC   fusion protein.
XX
SQ   Sequence 3939 BP; 1046 A; 1103 C; 1106 G; 684 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 3939;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2674 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2733

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2734 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2793

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2794 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2853

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     2854 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2913

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2914 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2973

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2974 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3033

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     3034 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3093

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     3094 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3153

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3154 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3213

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     3214 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3273

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3274 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3333
```

```
Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3334 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3393

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3394 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3453

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3454 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3513

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3514 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3573

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    3574 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3633

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3634 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3693

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3694 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3753

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    3754 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3813

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3814 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3873

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||||||||||||||||||||
Db    3874 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912


RESULT 12
AAS43902
ID   AAS43902 standard; DNA; 4170 BP.
XX
AC   AAS43902;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein 961c-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27605.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 23; Page 71-73; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4170 BP; 1185 A; 1138 C; 1119 G; 728 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 4170;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2905 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2964

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2965 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3024

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3025 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3084

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    3085 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3144

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3145 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3204

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3205 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3264

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    3265 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3324

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    3325 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3384

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3385 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3444

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    3445 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3504

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3505 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3564

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3565 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3624

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     3625 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3684

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3685 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3744

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3745 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3804

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     3805 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3864

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3865 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3924

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3925 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3984

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     3985 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4044

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4045 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4104

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     4105 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143


RESULT 13
AAD17055
ID   AAD17055 standard; DNA; 4170 BP.
XX
AC   AAD17055;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 961c-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta 961c-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4170
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 961c-983
FT                   fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10040.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 23; Page 31; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) 961c-983 fusion
CC   protein.
XX
SQ   Sequence 4170 BP; 1185 A; 1138 C; 1119 G; 728 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 4170;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2905 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2964

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2965 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3024

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3025 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3084

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     3085 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3144

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3145 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3204

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3205 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3264

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     3265 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3324

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     3325 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3384

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3385 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3444

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     3445 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3504

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3505 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3564

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3565 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3624

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3625 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3684
```

```
Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3685 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3744

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3745 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3804

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     3805 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 3864

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3865 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 3924

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3925 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3984

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     3985 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4044

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4045 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4104

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     4105 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143


RESULT 14
AAS43877
ID   AAS43877 standard; DNA; 4179 BP.
XX
AC   AAS43877;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-961c DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD . 07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27578.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 47-49; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 4179;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAGGCAGTGCT 2805

Qy     2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy     2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy     2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy     3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
            |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db     2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy     3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy     3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||
Db     3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144
```

RESULT 15
AAD17042
ID   AAD17042 standard; DNA; 4179 BP.
XX
AC   AAD17042;
XX
DT   29-NOV-2001 (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-961c fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G983-961c fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..4179
FT   /*tag= a
FT   /product= "N. meningitidis strain 2996 delta
FT   G983-961c fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10025.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state -
XX
PS   Example 3; Page 17-18; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-961c
CC   fusion protein.
XX
SQ   Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;

```
  Query Match             25.9%;  Score 831;  DB 22;  Length 4179;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy     2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy     2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy     2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy     2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy     2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy     2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy     2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
            |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db     2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy     2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy     2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy     2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy     2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy     2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy     2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy     2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAGGCAGTGCT 2862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||||||||||||||||||||
Db    3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144


RESULT 16
AAS43899
ID   AAS43899 standard; DNA; 4335 BP.
XX
AC   AAS43899;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein 961-983 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27602.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 23; Page 68-69; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 4335;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3070 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3129

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3130 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3189

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3190 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3249

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    3250 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3309

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3310 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3369

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3370 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3429

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    3430 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3489

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    3490 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3549

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3550 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3609

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    3610 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3669

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3670 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3729

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3730 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3789

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3790 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3849

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3850 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3909

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3910 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3969

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    3970 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 4029
```

```
Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4030 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 4089

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4090 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4149

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   4150 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4209

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4210 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4269

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||||||||||||||||||||
Db   4270 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308


RESULT 17
AAD17052
ID   AAD17052 standard; DNA; 4335 BP.
XX
AC   AAD17052;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 961-983 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein; 961-983 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4335
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 961-983
FT                   fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10037.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 23; Page 28-29; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) 961-983 fusion
CC   protein.
XX
SQ   Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 4335;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3070 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 3129

Qy   2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3130 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 3189

Qy   2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3190 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 3249

Qy   2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   3250 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 3309

Qy   2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3310 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 3369

Qy   2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3370 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 3429

Qy   2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   3430 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 3489

Qy   2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
          |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   3490 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 3549

Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3550 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 3609

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3610 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 3669

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3670 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 3729

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3730 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 3789

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3790 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 3849

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3850 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 3909

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3910 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 3969

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   3970 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 4029

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   4030 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 4089

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4090 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 4149

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   4150 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 4209

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4210 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 4269

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||||||||||||||||||||
Db   4270 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308
```

RESULT 18
AAS43876
ID   AAS43876 standard; DNA; 4344 BP.
XX
AC   AAS43876;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-961 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB, Comanducci M, Galeotti C, Masignani V, Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27577.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 46-47; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4344 BP; 1219 A; 1191 C; 1167 G; 767 T; 0 other;

```
  Query Match             25.9%;  Score 831;  DB 22;  Length 4344;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy   2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy   2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy   2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy   2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy   2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy   2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy   2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
          |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985
```

```
Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy    3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||||||||||||||||||||
Db    3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144


RESULT 19
AAS43874
ID   AAS43874 standard; DNA; 4425 BP.
XX
AC   AAS43874;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis fusion protein delta-G983-ORF46.1 DNA.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KW   Neisserial protein.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
DR   P-PSDB; AAU27575.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Example 15; Page 42-43; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
CC   and peptide regions of proteins of the invention.
XX
SQ   Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 4425;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy    2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy    2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy    2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy    2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy    2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy    2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy    2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy    2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy    2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy    2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy    2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy    2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy    2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy    2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy    2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy    2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy    2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy    3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
           |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db    2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy    3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105

Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||||||||||||||||||||
Db   3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144


RESULT 20
AAD17039
ID   AAD17039 standard; DNA; 4425 BP.
XX
AC   AAD17039;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-ORF46.1 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein; open reading frame; ORF;
KW   delta G983-ORF46.1 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4425
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G983-ORF46.1 fusion protein"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10022.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 3; Page 13-14; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-ORF46.1
CC   (open reading frame) fusion protein.
XX
SQ   Sequence 4425 BP; 1165 A; 1259 C; 1219 G; 782 T; 0 other;

  Query Match             25.9%;  Score 831;  DB 22;  Length 4425;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1231;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1963 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1906 CTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCC 1965

Qy   2023 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1966 TATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCC 2025

Qy   2083 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2026 GCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAA 2085

Qy   2143 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGC 2202
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   2086 CTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGC 2145

Qy   2203 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2146 ACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAG 2205

Qy   2263 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2206 CATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCC 2265

Qy   2323 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGG 2382
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   2266 GACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGG 2325

Qy   2383 TTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2442
          |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   2326 TTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACG 2385

Qy   2443 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2386 TGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC 2445

Qy   2503 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGG 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2446 GCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGG 2505

Qy   2563 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2506 AGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAC 2565

Qy   2623 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2566 GATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGC 2625

Qy   2683 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2626 ATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATG 2685

Qy   2743 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2686 CAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTC 2745

Qy   2803 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2746 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCT 2805

Qy   2863 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAG 2922
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2806 TTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAG 2865

Qy   2923 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2866 CTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGAC 2925

Qy   2983 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2926 CTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGC 2985

Qy   3043 AAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTC 3102
          |||||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||
Db   2986 AAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTC 3045

Qy   3103 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3046 GAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTAC 3105
```

```
Qy   3163 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||||||||||||||||||||
Db   3106 GGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3144


RESULT 21
AAD17041
ID   AAD17041 standard; DNA; 3798 BP.
XX
AC   AAD17041;
XX
DT   29-NOV-2001  (first entry)
XX
DE   N. meningitidis strain 2996 delta G983-961 fusion DNA.
XX
KW   Heterologous expression; Neisserial protein;
KW   delta G983-961 fusion protein; ds.
XX
OS   Neisseria meningitidis 2996.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3798
FT                   /*tag=  a
FT                   /product= "N. meningitidis strain 2996 delta
FT                   G983-961 fusion protein"
FT                   /transl_except= (pos:2104..2109, aa:Xaa)
FT                   /note= "Xaa corresponds to residues 702-885 of delta
FT                   G983-961 fusion protein; There is apparent deletion
FT                   of codons corresponding to residues 702-885 of this
FT                   protein, which alters the reading frame"
XX
PN   WO200164920-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00420.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Giuliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-557776/62.
DR   P-PSDB; AAE10024.
XX
PT   Heterologous expression for the expression of two or more Neisserial
PT   proteins in fused state  -
XX
PS   Example 3; Page 16-17; 52pp; English.
XX
CC   The present invention relates to a method for simultaneous heterologous
CC   expression of two or more Neisserial proteins which are in a fused
CC   state. The method is useful for simultaneous heterologous expression of
CC   two or more Neisserial proteins. A protein that may be unstable or
CC   poorly expressed on its own is assisted by adding a suitable hybrid
CC   partner and commercial manufacture is simplified-only one expression and
CC   purification need to be employed in order to produce two separately-
CC   useful proteins. The present sequence is a DNA encoding
CC   Neisseria meningitidis (serogroup B, strain 2996) delta G983-961
CC   fusion protein.
XX
SQ   Sequence 3798 BP; 1079 A; 1020 C; 1007 G; 692 T; 0 other;

  Query Match             12.4%;  Score 398;  DB 22;  Length 3798;
  Best Local Similarity   99.6%;  Pred. No. 5.8e-183;
  Matches  498;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   1168 TCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGA 1227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1111 TCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGA 1170

Qy   1228 ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCG 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1171 ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCG 1230

Qy   1288 TGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCA 1347
          |||||||||||||||||||||||||||||| ||||||| |||||||||||||||||||||
Db   1231 TGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCA 1290

Qy   1348 GTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGA 1407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1291 GTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGA 1350

Qy   1408 CCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCC 1467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1351 CCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCC 1410

Qy   1468 TACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAA 1527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411 TACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAA 1470

Qy   1528 CTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTG 1587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471 CTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTG 1530

Qy   1588 GTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTAT 1647
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1531 GTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTAT 1590

Qy   1648 AACGGGGCGGCATCCGGCGG 1667
          ||||||||||||||||||||
Db   1591 AACGGGGCGGCATCCGGCGG 1610


RESULT 22
AAZ53298
ID   AAZ53298 standard; DNA; 526 BP.
XX
AC   AAZ53298;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 139 partial DNA sequence SEQ ID NO:545.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
```

```
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74536.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 393; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 526 BP; 159 A; 142 C; 131 G; 94 T; 0 other;

  Query Match             5.9%;  Score 189;  DB 21;  Length 526;
  Best Local Similarity   99.6%;  Pred. No. 2.4e-81;
  Matches  239;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      285 CCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTT 344
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      287 CCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTT 346

Qy      345 GATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTAT 404
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      347 GATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTAT 406

Qy      405 CGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGA 464
            |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      407 CGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGA 466

Qy      465 ACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGA 524
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      467 ACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGA 526


RESULT 23
AAV03603
ID   AAV03603 standard; DNA; 286 BP.
XX
AC   AAV03603;
XX
DT   22-OCT-1998  (first entry)
XX
DE   Neisseria meningitidis DNA sequence E85.
XX
KW   N. gonorrheae; N. lactamica; chromosome Z2491; region 1; region 2;
KW   region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
KW   meningitis; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9802547-A2.
XX
PD   22-JAN-1998.
XX
PF   11-JUL-1997;   97WO-FR01295.
XX
PR   12-JUL-1996;   96FR-0008768.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SMIK ) SMITHKLINE BEECHAM.
XX
PI   Nassif X,  Tinsley C,  Achtman M,  Merker P,  Ruelle J;
PI   Vinals C;
XX
DR   WPI; 1998-110594/10.
XX
PT   Genes present in Neisseria meningitidis but not other Neisseria
PT   species - and related host cells, RNA, anti-sense sequences,
PT   polypeptide(s) and antibodies, useful for diagnosing Neisseria
PT   meningitidis infection and in protective vaccines
XX
PS   Example 4; Page 128; 150pp; French.
XX
CC   AAV03575-606 represent sequences that are present in Neisseria
CC   meningitidis and N. gonorrheae but not in N. lactamica, except for the
CC   genes involved in biosynthesis of the capsule polysaccharide, frpA or C,
CC   opc, porA, rotamase, sequence IC1106, IgA protease, pillin, pilC,
CC   proteins which bind transferrin and opacity proteins. The DNA sequences
CC   are responsible for the differences in pathogenicity between N.
CC   meningitidis and N. gonorrheae, specifically they include the genes that
CC   allow N. meningitidis to cross the blood-brain barrier. DNA sequences
CC   common to N. meningitidis and N. gonorrheae, but absent from N.
CC   lactamica, are responsible for colonisation and penetration of the
CC   mucosa. The DNA sequences can be used to produce probes and primers, and
CC   antibodies produced against the encoded proteins are used in standard
CC   hybridisation/immunoassay processes for diagnosis of N. meningitidis
CC   infection, particularly meningitis.
XX
SQ   Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

  Query Match             5.5%;  Score 177;  DB 19;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-75;
  Matches  177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy      925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy      985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 1041
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 177


RESULT 24
AAA15396
ID   AAA15396 standard; DNA; 286 BP.
XX
AC   AAA15396;
XX
DT   04-SEP-2000  (first entry)
XX
DE   Genomic fragment of Neisseria meningitidis Z2491.
XX
KW   Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO200026375-A2.
XX
PD   11-MAY-2000.
XX
PF   28-OCT-1999;   99WO-FR02643.
XX
PR   30-OCT-1998;   98FR-0013693.
XX
PA   (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Aujame L,  Bouchardon A,  Renauld-Mongenie G,  Rokbi B,  Nassif X;
PI   Tinsley C,  Perrin A;
```

```
XX
DR   WPI; 2000-365622/31.
XX
PT   New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT   or preventative vaccines and for diagnosis  -
XX
PS   Claim 1; Page 44; 187pp; French.
XX
CC   The present sequence represents a genomic fragment of Neisseria
CC   meningitidis. The specificatino describes proteins that are specific
CC   for pathogenic strains of Neisseria. The polynucleotides, polypeptides,
CC   or their antigenic fragments, are used in vaccines to treat or protect
CC   against Neisseria infections, particularly by N. meningitidis. The
CC   polynucleotide sequences are also used for recombinant production of
CC   the polypeptide and to produce attenuated Neisseria strains that
CC   overexpress it, or express it in a non-toxic mutant form.
CC   note: the present sequence is included in Claim 1, but it is those
CC   sequences that do not include the present sequence that are actually
CC   claimed.
XX
SQ   Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

  Query Match              5.5%;  Score 177;  DB 21;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-75;
  Matches  177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy    925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy    985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 1041
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACA 177


RESULT 25
AAZ53297
ID   AAZ53297 standard; DNA; 528 BP.
XX
AC   AAZ53297;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 138 partial DNA sequence SEQ ID NO:543.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74535.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 393; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 528 BP; 159 A; 142 C; 131 G; 96 T; 0 other;

  Query Match              5.0%;  Score 160;  DB 21;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 3e-67;
  Matches  160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    133 ACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCC 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 ACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCC 195

Qy    193 GGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTT 252
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 GGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTT 255

Qy    253 GCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCC 292
          ||||||||||||||||||||||||||||||||||||||||
Db    256 GCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCC 295


RESULT 26
AAZ53296
ID   AAZ53296 standard; DNA; 537 BP.
XX
AC   AAZ53296;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria gonorrheae ORF 138 partial DNA sequence SEQ ID NO:541.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria gonorrheae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
```

```
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74534.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 392; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 537 BP; 169 A; 136 C; 130 G; 102 T; 0 other;

  Query Match             4.7%;  Score 149;  DB 21;  Length 537;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-62;
  Matches  149;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    344 TGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTA 403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 TGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTA 403

Qy    404 TCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAG 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 TCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAG 463

Qy    464 AACACGGCTATAACGAAAATTACAAAAAC 492
          |||||||||||||||||||||||||||||
Db    464 AACACGGCTATAACGAAAATTACAAAAAC 492


RESULT 27
AAZ53299
ID   AAZ53299 standard; DNA; 1365 BP.
XX
AC   AAZ53299;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria gonorrheae ORF 140 partial DNA sequence SEQ ID NO:547.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria gonorrheae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY74537.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 394; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1365 BP; 324 A; 405 C; 416 G; 220 T; 0 other;

  Query Match             4.1%;  Score 131;  DB 21;  Length 1365;
  Best Local Similarity   99.5%;  Pred. No. 3.6e-53;
  Matches  181;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1960 GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTG 2019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    121 GGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCCG 180

Qy   2020 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 2079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCG 240

Qy   2080 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 2139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTC 300

Qy   2140 GA 2141
          ||
Db    301 GA 302


RESULT 28
AAA81823/c
ID   AAA81823 standard; DNA; 476 BP.
XX
AC   AAA81823;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_370 SEQ ID NO:370.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
```

```
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1608; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 476 BP; 84 A; 154 C; 125 G; 111 T; 2 other;

  Query Match              1.8%;  Score 57;  DB 21;  Length 476;
  Best Local Similarity  100.0%;  Pred. No. 3.5e-17;
  Matches   57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1072 GCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAT 1128
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      467 GCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAT 411


RESULT 29
AAA82003
ID   AAA82003 standard; DNA; 577 BP.
XX
AC   AAA82003;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_550 SEQ ID NO:550.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1652-1653; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 577 BP; 189 A; 134 C; 99 G; 155 T; 0 other;

  Query Match              1.3%;  Score 41;  DB 21;  Length 577;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-09;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       44 CTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGC 84
            |||||||||||||||||||||||||||||||||||||||||
Db      537 CTGCCATGGCGTTAGCTGTTGCAACAACACTTTCTGCCTGC 577


RESULT 30
AAS44248
ID   AAS44248 standard; DNA; 50 BP.
XX
AC   AAS44248;
XX
```

```
DT   18-DEC-2001  (first entry)
XX
DE   Neisseria meningitidis B MC58 genomic DNA sequence PCR primer #343.
XX
KW   Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ss;
KW   Neisserial protein; PCR primer.
XX
OS   Neisseria meningitidis.
OS   Synthetic.
XX
PN   WO200164922-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-IB00452.
XX
PR   28-FEB-2000; 2000GB-0004695.
PR   13-NOV-2000; 2000GB-0027675.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Arico MB,  Comanducci M,  Galeotti C,  Masignani V,  Guiliani MM;
PI   Pizza M;
XX
DR   WPI; 2001-582163/65.
XX
PT   Producing heterologous proteins from Neisseria meningitidis and N.
PT   gonorrhoeae -
XX
PS   Disclosure; Page 89; 119pp; English.
XX
CC   The invention relates to methods for the heterologous expression of
CC   Neisserial proteins from Neisseria meningitidis and Neisseria
CC   gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC   leader peptide, and may be replaced by a domain from a different protein
CC   to make a fusion protein, in order to enhance heterologous expression of
CC   Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC   stretch, can be mutated to enhance expression. The proteins used in the
CC   processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC   AAS43807-AAS43867 and AAS43906-AAS44358 represent PCR primers used in the
CC   methods of the invention.
XX
SQ   Sequence 50 BP; 15 A; 20 C; 7 G; 8 T; 0 other;

  Query Match             1.1%;  Score 35;  DB 22;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-06;
  Matches   35;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 CGAACGACCCCAACCTTCCCTACAAAAACTTTCAA 38
            |||||||||||||||||||||||||||||||||||
Db     16 CGAACGACCCCAACCTTCCCTACAAAAACTTTCAA 50


RESULT 31
ABL05527
ID   ABL05527 standard; cDNA; 2280 BP.
XX
AC   ABL05527;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 11063.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB61424.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 11063; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2280 BP; 518 A; 605 C; 561 G; 596 T; 0 other;

  Query Match             0.8%;  Score 25;  DB 23;  Length 2280;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db     17 GCAGCAACAGCAGAGCAACAACAGC 41


RESULT 32
ABL28074/c
ID   ABL28074 standard; DNA; 3302 BP.
XX
AC   ABL28074;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 35695.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
```

```
XX
PS   Claim 1; SEQ ID NO 35695; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3302 BP; 765 A; 737 C; 817 G; 983 T; 0 other;

  Query Match             0.8%;  Score 25;  DB 23;  Length 3302;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db    231 GCAGCAACAGCAGAGCAACAACAGC 207


RESULT 33
ABL05526/c
ID   ABL05526 standard; cDNA; 13327 BP.
XX
AC   ABL05526;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 11060.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB61423.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 11060; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13327 BP; 3971 A; 2726 C; 2711 G; 3919 T; 0 other;

  Query Match             0.8%;  Score 25;  DB 23;  Length 13327;
  Best Local Similarity   100.0%;  Pred. No. 0.11;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    143 GCAGCAACAGCAGAGCAACAACAGC 167
          |||||||||||||||||||||||||
Db  12311 GCAGCAACAGCAGAGCAACAACAGC 12287


RESULT 34
AAA15341/c
ID   AAA15341 standard; DNA; 30 BP.
XX
AC   AAA15341;
XX
DT   04-SEP-2000  (first entry)
XX
DE   PCR primer for a Neisseria pathogenic strain DNA sequence.
XX
KW   Pathogenic strain; Neisseria; vaccine; Neisseria infection;
KW   PCR primer; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO200026375-A2.
XX
PD   11-MAY-2000.
XX
PF   28-OCT-1999;   99WO-FR02643.
XX
PR   30-OCT-1998;   98FR-0013693.
XX
PA   (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Aujame L,  Bouchardon A,  Renauld-Mongenie G,  Rokbi B,  Nassif X;
PI   Tinsley C,  Perrin A;
XX
DR   WPI; 2000-365622/31.
XX
PT   New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT   or preventative vaccines and for diagnosis  -
XX
PS   Example 1; Page 27; 187pp; French.
XX
CC   The present sequence represents a PCR primer that was used to
CC   amplify DNA encoding a protein that is specific for pathogenic
CC   strains of Neisseria. The polynucleotides, polypeptides, or their
CC   antigenic fragments, are used in vaccines to treat or protect against
CC   Neisseria infections, particularly by N. meningitidis. The
CC   polynucleotide sequence is also used for recombinant production of
CC   the polypeptide and to produce attenuated Neisseria strains that
CC   overexpress it, or express it in a non-toxic mutant form.
XX
SQ   Sequence 30 BP; 7 A; 11 C; 9 G; 3 T; 0 other;

  Query Match             0.7%;  Score 23;  DB 21;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   3181 GTCGGCGTAGGCTACCGGTTCTG 3203
          |||||||||||||||||||||||
Db     30 GTCGGCGTAGGCTACCGGTTCTG 8


RESULT 35
ABL27930
ID   ABL27930 standard; DNA; 2382 BP.
XX
```

```
AC   ABL27930;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 35263.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 35263; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2382 BP; 685 A; 525 C; 530 G; 642 T; 0 other;

  Query Match             0.7%;  Score 22;  DB 23;  Length 2382;
  Best Local Similarity   100.0%;  Pred. No. 3.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    266 GGGATGCCAAAATCAATGCCCC 287
          ||||||||||||||||||||||
Db   2340 GGGATGCCAAAATCAATGCCCC 2361


RESULT 36
AAA15340
ID   AAA15340 standard; DNA; 34 BP.
XX
AC   AAA15340;
XX
DT   04-SEP-2000  (first entry)
XX
DE   PCR primer for a Neisseria pathogenic strain DNA sequence.
XX
KW   Pathogenic strain; Neisseria; vaccine; Neisseria infection;
KW   PCR primer; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO200026375-A2.
XX
PD   11-MAY-2000.
XX
PF   28-OCT-1999;   99WO-FR02643.
XX
PR   30-OCT-1998;   98FR-0013693.
XX
PA   (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Aujame L,  Bouchardon A,  Renauld-Mongenie G,  Rokbi B,  Nassif X;
PI   Tinsley C,  Perrin A;
XX
DR   WPI; 2000-365622/31.
XX
PT   New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT   or preventative vaccines and for diagnosis  -
XX
PS   Example 1; Page 27; 187pp; French.
XX
CC   The present sequence represents a PCR primer that was used to
CC   amplify DNA encoding a protein that is specific for pathogenic
CC   strains of Neisseria. The polynucleotides, polypeptides, or their
CC   antigenic fragments, are used in vaccines to treat or protect against
CC   Neisseria infections, particularly by N. meningitidis. The
CC   polynucleotide sequence is also used for recombinant production of
CC   the polypeptide and to produce attenuated Neisseria strains that
CC   overexpress it, or express it in a non-toxic mutant form.
XX
SQ   Sequence 34 BP; 9 A; 15 C; 5 G; 5 T; 0 other;

  Query Match             0.7%;  Score 21;  DB 21;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGCGAACGACCCCAACCTTC 21
          |||||||||||||||||||||
Db     14 ATGCGAACGACCCCAACCTTC 34


RESULT 37
AAZ55456
ID   AAZ55456 standard; DNA; 34 BP.
XX
AC   AAZ55456;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria species ORF cloning PCR primer #841.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; PCR primer; ss.
XX
OS   Synthetic.
OS   Neisseria sp.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
```

```
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Example 16; Page 164; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 34 BP; 9 A; 10 C; 7 G; 8 T; 0 other;

  Query Match              0.6%;  Score 20;  DB 21;  Length 34;
  Best Local Similarity  100.0%;  Pred. No. 36;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      54 GTTAGCTGTTGCAACAACAC 73
           ||||||||||||||||||||
Db      15 GTTAGCTGTTGCAACAACAC 34


RESULT 38
AAT45009/c
ID   AAT45009 standard; DNA; 735 BP.
XX
AC   AAT45009;
XX
DT   13-AUG-1997  (first entry)
XX
DE   Serratia marcescens IAM 13540 acid phosphatase DNA.
XX
KW   IAM 13540; acid phosphatase; production; nucleoside; 5'-phosphate;
KW   ester; condiment; pharmaceutical; intermediate; ds.
XX
OS   Serratia marcescens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..735
FT                   /*tag=  a
FT                   /product= acid_phosphatase
XX
PN   WO9637603-A1.
XX
PD   28-NOV-1996.
XX
PF   24-MAY-1996;   96WO-JP01402.
XX
PR   26-MAR-1996;   96JP-0094680.
PR   25-MAY-1995;   95JP-0149781.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Asano Y,  Mihara Y,  Utagawa T,  Yamada H;
XX
DR   WPI; 1997-021215/02.
DR   P-PSDB; AAW06460.
XX
PT   Efficient production of nucleoside 5'-phosphate - by reaction of a
PT   nucleoside with a phosphoric acid donor in the presence of an acid
PT   phosphatase
XX
PS   Example 24; Pages 74-75; 95pp; Japanese.
XX
CC   The present sequence encodes the Serratia marcescens IAM 13540
CC   acid phosphatase (AP), which can be used to produce a nucleoside
CC   5'-phosphate ester from the corresponding nucleoside when a
CC   phosphate donor, e.g. poly-, phenyl- or carbamyl-phosphoric acid,
CC   is reacted in its presence at pH 3.0 to 5.5. The PA can be used for
CC   the economic and efficient production of nucleoside 5'-phosphate
CC   esters for use as condiments, pharmaceuticals and intermediates for
CC   pharmaceuticals.
XX
SQ   Sequence 735 BP; 180 A; 215 C; 198 G; 142 T; 0 other;

  Query Match              0.6%;  Score 20;  DB 18;  Length 735;
  Best Local Similarity  100.0%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1751 TGCAGCTGGGCGGCGAAGGT 1770
           ||||||||||||||||||||
Db     698 TGCAGCTGGGCGGCGAAGGT 679


RESULT 39
AAV43062/c
ID   AAV43062 standard; DNA; 735 BP.
XX
AC   AAV43062;
XX
DT   21-OCT-1998  (first entry)
XX
DE   DNA encoding an acid phosphatase enzyme.
XX
KW   Acid phosphatase; preparation; nucleoside 5'-phosphate ester;
KW   seasoning; intermediate; ds.
XX
OS   Serratia ficaria.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..735
FT                   /*tag=  a
XX
PN   EP857788-A2.
XX
PD   12-AUG-1998.
XX
PF   20-NOV-1997;   97EP-0309365.
XX
PR   18-JUN-1997;   97JP-0161674.
PR   21-NOV-1996;   96JP-0311103.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Asano Y,  Mihara Y,  Utagawa T,  Yamada H;
XX
DR   WPI; 1998-416010/36.
DR   P-PSDB; AAW71033.
XX
PT   Preparation of nucleoside 5'-phosphates comprises reacting
PT   nucleoside with phosphate donor in presence of acid phosphatase -
PT   used as seasonings or pharmaceutical intermediates
XX
PS   Example 24; Pages 47-48; 83pp; English.
XX
CC   The present sequence encodes an acid phosphatase enzyme. The
CC   specification describes a method for the preparation of nucleoside
CC   5'-phosphate esters. The method comprises reacting a nucleoside with
CC   a phosphate donor at pH 3.0-5.5 in the presence of an acid phosphatase
CC   that has been altered to increase its affinity for the nucleoside and/or
CC   to increase its thermal stability, or in the presence of a microorganism
CC   that has been transformed with recombinant DNA containing a gene coding
```

```
CC   for such an acid phosphatase. Nucleoside 5'-phosphates are useful as
CC   seasonings or pharmaceuticals or as intermediates for them.
CC   note: this sequence appears to be claimed (claim 6), but as the claim
CC   refers to amino acid sequences, it is clear that the corressponding
CC   protein is being claimed.
XX
SQ   Sequence 735 BP; 180 A; 215 C; 198 G; 142 T; 0 other;

  Query Match             0.6%;  Score 20;  DB 19;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1751 TGCAGCTGGGCGGCGAAGGT 1770
          ||||||||||||||||||||
Db    698 TGCAGCTGGGCGGCGAAGGT 679


RESULT 40
ABQ32992/c
ID   ABQ32992 standard; DNA; 1105 BP.
XX
AC   ABQ32992;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 19583.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP10074.
XX
PR   01-SEP-2000; 2000DE-1043826.
PR   05-SEP-2000; 2000DE-1044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful
PT   for diagnosis and prognosis, comprises selective hybridization of
PT   amplicons from chemically treated DNA  -
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one
CC   member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC   and the degree of hybridisation to both classes is determined from the
CC   label on the amplicon. From the ratio of labels hybridised to the two
CC   classes of oligomers, the degree of methylation is calculated. The method
CC   is used: (i) for diagnosis and/or prognosis of side effects of
CC   therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC   of the central nervous, cardiovascular, gastrointestinal and respiratory
CC   systems etc., particularly by detecting mutations or single nucleotide
CC   polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC   types and for investigating cell differentiation. The method allows the
CC   methylation status of many C residues to be determined simultaneously.
CC   ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC   method for determining the degree of cytosine methylation described in
CC   the disclosure of the invention.
XX
SQ   Sequence 1105 BP; 144 A; 147 C; 416 G; 398 T; 0 other;

  Query Match             0.6%;  Score 20;  DB 24;  Length 1105;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    475 AACGAAAATTACAAAAACTA 494
          ||||||||||||||||||||
Db    353 AACGAAAATTACAAAAACTA 334


Search completed: January 27, 2003, 20:57:25
Job time : 1443 secs
```

OM nucleic - nucleic search, using sw model

Run on: January 27, 2003, 16:40:01 ; Search time 5468 Seconds
(without alignments)
17052.924 Million cell updates/sec

Title: US-09-830-433A-7
Perfect score: 3204
Sequence: 1 atgcgaacgaccccaacctt..........gcgtaggctaccggttctga 3204

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3204	100.0	3204	6	AX024064	AX024064 Sequence
	2	3202.4	100.0	3204	1	NME311654	AJ311654 Neisseria
c	3	3160.6	98.6	349061	1	NMA2Z2491	AL162753 Neisseria
	4	3062.2	95.6	10591	1	AE002545	AE002545 Neisseria
	5	3062.2	95.6	349980	6	AX044034	AX044034 Sequence
	6	3042.4	95.0	3254	1	NME277537	AJ277537 Neisseria
	7	2986.2	93.2	4218	6	AX236469	AX236469 Sequence
	8	2986.2	93.2	4218	6	AX239755	AX239755 Sequence
	9	2983.8	93.1	4335	6	AX236457	AX236457 Sequence
	10	2983.8	93.1	4335	6	AX239743	AX239743 Sequence
	11	2983.2	93.1	3939	6	AX236419	AX236419 Sequence
	12	2983.2	93.1	3939	6	AX239729	AX239729 Sequence
	13	2983.2	93.1	4170	6	AX236463	AX236463 Sequence
	14	2983.2	93.1	4170	6	AX239749	AX239749 Sequence
	15	2971.8	92.8	3939	6	AX236409	AX236409 Sequence
	16	2971.8	92.8	3939	6	AX239717	AX239717 Sequence
	17	2971.8	92.8	4179	6	AX236413	AX236413 Sequence
	18	2971.8	92.8	4179	6	AX239721	AX239721 Sequence
	19	2971.8	92.8	4344	6	AX236411	AX236411 Sequence
	20	2971.8	92.8	4344	6	AX239719	AX239719 Sequence
	21	2971.8	92.8	4425	6	AX236407	AX236407 Sequence
	22	2971.8	92.8	4425	6	AX239715	AX239715 Sequence
c	23	576.4	18.0	600	1	AF169448	AF169448 Neisseria
	24	270.8	8.5	286	1	AF169473	AF169473 Neisseria
	25	270.8	8.5	286	6	A68924	A68924 Sequence 95
	26	147.8	4.6	3493	1	BPE318229	AJ318229 Bordetell
	27	132.4	4.1	11292	1	AE011850	AE011850 Xanthomon
c	28	130.2	4.1	11019	1	AE012308	AE012308 Xanthomon
	29	81.6	2.5	1434	1	NGOPAK	X52364 N.gonorrhoe
	30	77.4	2.4	10807	1	AE011766	AE011766 Xanthomon
	31	77	2.4	10029	1	AE012229	AE012229 Xanthomon
	32	72.8	2.3	10495	1	AE003880	AE003880 Xylella f
	33	72.8	2.3	14988	1	AB015053	AB015053 Pseudomon
	34	71.6	2.2	12152	1	AF286062	AF286062 Pseudomon
	35	71.2	2.2	11186	1	AE003939	AE003939 Xylella f
	36	70.8	2.2	215050	1	AL646057	AL646057 Ralstonia
	37	68	2.1	7030	1	SMASSPH	D78380 Serratia ma
	38	67	2.1	11476	1	AE004006	AE004006 Xylella f
	39	65	2.0	5423	1	AF216702	AF216702 Pseudomon
	40	62	1.9	10732	6	E32986	E32986 Gene encodi
c	41	60.2	1.9	67200	1	MTV017	AL021897 Mycobacte
c	42	56	1.7	125020	9	AF429315	AF429315 Homo sapi
	43	53.2	1.7	125020	9	AF429315	AF429315 Homo sapi
	44	53	1.7	3510	1	PTO7828	AJ007828 Pseudomon
c	45	53	1.7	10165	1	AE001935	AE001935 Deinococc

ALIGNMENTS

RESULT 1
AX024064
LOCUS AX024064 3204 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 7 from Patent FR2785293.
ACCESSION AX024064
VERSION AX024064.1 GI:10184376
KEYWORDS .
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 3204)
AUTHORS Nassif,X., Tinsley,C., Aujame,L., Perrin,A., Rokbi,B.,
Bouchardon,A. and Renauld,M.G.
JOURNAL Patent: FR 2785293-A 7 05-MAY-2000;

```
            PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES             Location/Qualifiers
    source           1..3204
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
    CDS              1..3204
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC08908.1"
                     /db_xref="GI:10184377"
                     /translation="MRTTPTFPTKTFKPAAMALAVATTLSACLGGGGGTSAPDFNAGG
                     TGIGSNSRATTAKSAAVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPPNLHT
                     GDFTNPNDAYKNLINLKPAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNE
                     NYKNYTAYMRKEAPEDGGGKDIKASFDDEAVIETEAKPTDIRHVKEIGHIDVVSHIIG
                     GRSVDGRPAGGIAPDATLHIMNTHDGTKNEIMSAAIRNAWVKLGERGVRIVNNSFGTT
                     SRAGTADHFQIANSEEQYRQALLAYSGGDKTDEGIRLMQQSDYGNLSYHIRNKNMLFI
                     FSASNDAQAQPNTLTLLPFYEKDAQKGIITVAGVDRSGEKFNGSNHCGITAMWCLSAP
                     YEASVRFTRTNPIQIAGTSFSAPIVTGTAALLLQKYPWMSNDNLRTTLLTTAQDIGAV
                     GVDSKFGWGLLDAGKAMNGPASFPFGDFTADTKGTSDIAYSFRNDISGTGGLIKKGGS
                     QLQLHGNNTYTGKTIIEGGSLVLYGNNKSDMRVETKGALIYNGAASGGSLNSDGIVYL
                     ADTDRSGANETVHIKGDLQLGGEGTLYTRLGKLLKVDGTAMTGGKLYMSARGKGAGYL
                     NRTGQRVPFLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGDTLSYYVRRGNA
                     ARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAADRTDMPGI
                     RPYGATFRAAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLKAVSDGLDHN
                     ATGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHSTWSE
                     NSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTLM
                     QLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLTEGTLVGLAG
                     LKLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGL
                     GADVEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRF"
BASE COUNT      831 a    909 c    899 g    565 t
ORIGIN

  Query Match            100.0%;  Score 3204;  DB 6;  Length 3204;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120

Qy        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180

Qy        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240

Qy        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300

Qy        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360

Qy        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420

Qy        421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480

Qy        481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540

Qy        541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600

Qy        601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660

Qy        661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720

Qy        721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780

Qy        781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGAGGGCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGAGGGCAGGC 840

Qy        841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900

Qy        901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960

Qy        961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020

Qy       1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080

Qy       1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140

Qy       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440

Qy       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500

Qy       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560

Qy       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620

Qy       1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680

Qy    1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740

Qy    1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800

Qy    1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860

Qy    1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920

Qy    1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980

Qy    1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040

Qy    2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100

Qy    2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160

Qy    2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220

Qy    2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280

Qy    2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340

Qy    2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400

Qy    2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460

Qy    2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520

Qy    2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580

Qy    2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640

Qy    2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700

Qy    2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760

Qy    2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820

Qy    2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880

Qy    2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940

Qy    2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000

Qy    3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060

Qy    3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120

Qy    3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180

Qy    3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
             ||||||||||||||||||||||||
Db    3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
```

```
RESULT 2
NME311654
LOCUS       NME311654               3204 bp    DNA     linear   BCT 24-JUL-2002
DEFINITION  Neisseria meningitidis ausP gene for autotransporter serine
            protease.
ACCESSION   AJ311654
VERSION     AJ311654.1  GI:13397949
KEYWORDS    ausP gene; autotransporter serine protease.
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1
  AUTHORS   Turner,D.P., Wooldridge,K.G. and Ala'Aldeen,D.A.
  TITLE     Autotransported Serine Protease A of Neisseria meningitidis: an
            Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein
  JOURNAL   Infect. Immun. 70 (8), 4447-4461 (2002)
  MEDLINE   22112893
REFERENCE   2  (bases 1 to 3204)
  AUTHORS   Ala'Aldeen,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2001) Ala'Aldeen D.A., Microbiology, University
            of Nottingham, University Hospital, Nottingham, NG7 2UH, UNITED
            KINGDOM  .
COMMENT     Related sequence: AJ277537.
FEATURES             Location/Qualifiers
     source          1..3204
                     /organism="Neisseria meningitidis"
                     /strain="Group C ET37"
                     /isolate="Z4181"
                     /db_xref="taxon:487"
                     /country="Mali"
     gene            1..3204
                     /gene="ausP"
     CDS             1..3204
                     /gene="ausP"
                     /function="unknown"
                     /codon_start=1
```

```
                    /transl_table=11
                    /product="autotransporter serine protease (AusP)"
                    /protein_id="CAC34605.1"
                    /db_xref="GI:13397950"
                    /db_xref="SPTREMBL:Q9AE78"
                    /translation="MRTTPTFPTKTFKPAAMALAVATTLSACLGGGGGTSAPDFNAGG
                    TGIGSNSRATTAKSAAVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPPNLHT
                    GDFTNPNDAYKNLINLKPAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNE
                    NYKNYTAYMRKEAPEDGGGKDIKASFDDEAVIETEAKPTDIRHVKEIGHIDVVSHIIG
                    GRSVDGRPAGGIAPDATLHIMNTHDGTKNEIMSAAIRNAWVKLGERGVRIVNNSFGTT
                    SRAGTADHFQIANSEEQYRQALLAYSGGDKTDEGIRLMQQSDYGNLSYHIRNKNMLFI
                    FSASNDAQAQPNTLALLPFYEKDAQKGIITVAGVDRSGEKFNGSNHCGITAMWCLSAP
                    YEASVRFTRTNPIQIAGTSFSAPIVTGTAALLLQKYPWMSNDNLRTTLLTTAQDIGAV
                    GVDSKFGWGLLDAGKAMNGPASFPFGDFTADTKGTSDIAYSFRNDISGTGGLIKKGGS
                    QLQLHGNNTYTGKTIIEGGSLVLYGNNKSDMRVETKGALIYNGAASGGSLNSDGIVYL
                    ADTDRSGANETVHIKGDLQLGGEGTLYTRLGKLLKVDGTAMTGGKLYMSARGKGAGYL
                    NRTGQRVPFLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGDTLSYYVRRGNA
                    ARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAADRTDMPGI
                    RPYGATFRAAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLKAVSDGLDHN
                    ATGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHSTWSE
                    NSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTLM
                    QLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLTEGTLVGLAG
                    LKLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGL
                    GADVEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRF"
BASE COUNT      830 a    909 c    900 g    565 t
ORIGIN

  Query Match          100.0%;  Score 3202.4;  DB 1;  Length 3204;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3203;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTC 120

Qy        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCA 180

Qy        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGT 240

Qy        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTG 300

Qy        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT 360

Qy        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAA 420

Qy        421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAA 480

Qy        481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAA 540

Qy        541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGAT 600

Qy        601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGT 660

Qy        661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAAT 720

Qy        721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAG 780

Qy        781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGC 840

Qy        841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCC 900

Qy        901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGC 960

Qy        961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGAC 1020

Qy       1021 GCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       1021 GCACAAGCTCAGCCCAACACACTGGCCCTATTGCCATTTTATGAAAAAGATGCTCAAAAA 1080

Qy       1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCAT 1140

Qy       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440

Qy       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500

Qy       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560

Qy       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620

Qy       1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680

Qy       1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
```

```
Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180

Qy   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
          ||||||||||||||||||||||||
Db   3181 GTCGGCGTAGGCTACCGGTTCTGA 3204
```

```
RESULT 3
NMA2Z2491/c
LOCUS       NMA2Z2491              349061 bp    DNA     linear   BCT 04-DEC-2000
DEFINITION  Neisseria meningitidis serogroup A strain Z2491 complete genome;
            segment 2/7.
ACCESSION   AL162753 AL157959
VERSION     AL162753.2  GI:7379120
KEYWORDS    .
SOURCE      Neisseria meningitidis Z2491.
  ORGANISM  Neisseria meningitidis Z2491
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 349061)
  AUTHORS   Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
            Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
            Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
            Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
            Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
  TITLE     Complete DNA sequence of a serogroup A strain of Neisseria
            meningitidis Z2491
  JOURNAL   Nature 404 (6777), 502-506 (2000)
  MEDLINE   20222556
   PUBMED   10761919
REFERENCE   2  (bases 1 to 349061)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
            sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of N. meningitidis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES             Location/Qualifiers
     source          1..349061
                     /organism="Neisseria meningitidis Z2491"
                     /strain="Z2491"
                     /db_xref="taxon:122587"
                     /note="serogroup: A"
     repeat_unit     complement(24..206)
                     /note="ATR repeat; hmmfs hit to HMM ATR (1 - 183), score:
                     310.39"
                     /label=ATR
```

```
RBS             209. .212
gene            220. .681
                /gene="NMA0368"
CDS             220. .681
                /gene="NMA0368"
                /note="NMA0368, probable integral membrane protein, len:
                153 aa; contains four probable transmembrane domains"
                /codon_start=1
                /transl_table=11
                /product="putative integral membrane protein"
                /protein_id="CAB83669.1"
                /db_xref="GI:7379121"
                /db_xref="SPTREMBL:Q9JWH7"
                /translation="MQEQNRKSSFPIVMLLVSVALWIASLSNVAFYLGNHGSMEGLTV
                LILGSIFASLDIRYCAVYANYVWLAAIVLLALRKKVVPVHAAFWGLALVAFSVKAVYV
                DEAGNTSDIVRYGAGFYLWYAAFAVASIGTFAGKNKERKAASAADGTKNDV"
gene            671. .1492
                /gene="hemK"
CDS             671. .1492
                /gene="hemK"
                /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
                to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
                scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
                PS00092 N-6 Adenine-specific DNA methylases signature"
                /codon_start=1
                /transl_table=11
                /product="HemK protein"
                /protein_id="CAB83670.1"
                /db_xref="GI:7379122"
                /db_xref="SPTREMBL:Q9JWH6"
                /translation="MTFDKWLGLSKLPKNEARMLLQYVSEYTRVQLLTRGGEEMPDEV
                RQRADRLAQRRLNGEPVAYILGAREFYGRRFTVNPSVLIPRPETEHLVEAVLARLPEN
                GRVWDLGTGSGAVAVTVALERPDAFVRASDISPPALETARKNAADLGARVEFAYGSWF
                DTDMPSEGKWDIIVSNPPYIENGDKHLSQGDLRFEPQIALTDFSDGLSCIRTLAQGAP
                DRLAEGGFLLLEHGFDQGAAVRGVLAENGFSGVETLPDLAGLDRVTLGKYMKHLK"
misc_feature    831. .840
                /gene="hemK"
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    1162. .1171
                /gene="hemK"
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    1187. .1207
                /gene="hemK"
                /note="PS00092 N-6 Adenine-specific DNA methylases
                signature"
misc_feature    complement(1279. .1288)
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
RBS             1565. .1568
gene            1578. .2966
                /gene="NMA0370"
CDS             1578. .2966
                /gene="NMA0370"
                /note="NMA0370, probable integral membrane protein, len:
                462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
                P44640 hypothetical protein HI0325 (450 aa), fasta scores;
                E(): 0, 49.8% identity in 464 aa overlap"
                /codon_start=1
                /transl_table=11
                /product="putative integral membrane protein"
                /protein_id="CAB83671.1"
                /db_xref="GI:7379123"
                /db_xref="SPTREMBL:Q9JWH5"
                /translation="MNAVVVAVIVMLVLSLSRVHVVLSLTIGAFVGGAVAGMPLQNIA
                DAAGQVSQAGIIPVFNKGLEGGAKIALSYAMLGAFAMAITHSGLPQQLAGAVVRKLNR
                GGMPDSVRSGEGVVKWLLLSIILVMGMMSQNIIPIHIAFIPMIVPPLLLVFNRLKIDR
                RLIACVITFGLVTTYMFLPYGFGAIFLNEILLGNIHSAAPQLDVKNINVMAAMAIPAL
                GMLAGLLLAFVHYRKPRLYQSNNADTAGNADAANRPQPSAYRSLAAAVAIAVCFAIQL
                MYEDSLVLGAMLGFAVFMMLGVINRDKANDVFGEGIKMMAMVGFIMIAAQGFAAVMNA
                TGHIQPLVESSMAIFGNSKGMAALAMLVVGLLVTMGIGSSFSTLPIIAAIYVPLCVGL
                GFSPLATVAIVGTAGALGDAGSPASDSTLGPTMGLNADGQHDHIRDSVIPTFIHYNIP
                LLIAGWIAAMVL"
gene            2967. .3191
                /gene="slyX"
CDS             2967. .3191
                /gene="slyX"
                /note="NMA0371, slyX, SLYX protein homolog, len: 74 aa;
                simlar to SLYX_HAEIN P44759 SLYX protein homolog (73 aa),
                fasta scores; E(): 0.0017, 33.8% identity in 74 aa
                overlap, and SLYX_ECOLI P30857 SLYX protein (72 aa), fasta
                scores; E(): 0.24, 32.4% identity in 68 aa overlap"
                /codon_start=1
                /transl_table=11
                /product="SLYX protein homolog"
                /protein_id="CAB83672.1"
                /db_xref="GI:7379124"
                /db_xref="SPTREMBL:Q9JWH4"
                /translation="MDAVQEFEHRITELEIQSALQEDVIAGLNAMVAELRQTLDLQQA
                QLRLLYQKMQDRNPDAQEPYSLRDEIPPHY"
gene            complement(3271. .3423)
                /gene="NMA0372"
CDS             complement(3271. .3423)
                /gene="NMA0372"
                /note="NMA0372, unknown, questionable CDS, len: 50 aa"
                /codon_start=1
                /transl_table=11
                /product="very hypothetical protein NMA0372"
                /protein_id="CAB83673.1"
                /db_xref="GI:7379125"
                /db_xref="SPTREMBL:Q9JWH3"
                /translation="MFGRLSPLGSDGFFLAVYEYPVGFICLQGKTAYFRSGGKRFHRI
                KGILSD"
misc_feature    3562. .3571
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
gene            complement(3586. .4356)
                /gene="thiF"
CDS             complement(3586. .4356)
                /gene="thiF"
                /note="NMA0373, thiF, probable ThiF protein, len: 256 aa;
                similar to e.g. THIF_ECOLI P30138 THIF protein (251 aa),
                fasta scores; E(): 0, 43.1% identity in 246 aa overlap,
                and MOEB_ECOLI P12282 molybdopterin biosynthesis MOEB
                protein. (249 aa), fasta scores; E(): 0, 43.9% identity in
                244 aa overlap (note that N.m. does not have orthologs of
                any other molybdopterin biosynthesis proteins). Contains
                Pfam match to entry PF00899 ThiF_family, ThiF family"
                /codon_start=1
                /transl_table=11
                /product="ThiF protein"
                /protein_id="CAB83674.1"
                /db_xref="GI:7379126"
                /db_xref="SPTREMBL:Q9JWH2"
                /translation="MTTTEHDNDDAFLLRYSRHILLDEIGIEGQQKLSAAHILVVGCG
                GLGAAALPYLAASGIGTLTIADSDTVELHNLQRQVAFDEGDVGKLKTEALADRLRHIN
                HTVDVRTINEKLDGCRLTGLVQAADIVLDCCDNYATRQAVNRACVQAKTPLVSGAAVR
                FEGQLAVYRPDLPDSPCYACLFDGGSASDGICSLFGVFSPLVGIIGSTQAAEALKILL
                DAGEPSHGRLAVYRALEGGWQYFDLPRNPECPVCGAER"
misc_feature    3787. .3796
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    complement(3847. .4257)
                /gene="thiF"
                /note="Pfam match to entry PF00899 ThiF_family, ThiF
                family, score 186.60, E-value 4.1e-52"
misc_feature    complement(4418. .4427)
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
gene            4471. .7224
                /gene="ppc"
CDS             4471. .7224
                /gene="ppc"
                /EC_number="4.1.1.31"
                /note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
```

```
                  917 aa; simlar to many e.g. CAPP_RHOPA O32483
                  phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),
                  fasta scores; E(): 0, 43.3% identity in 928 aa overlap.
                  Contains 2x Pfam match to entry PF00311 PEPcase,
                  Phosphoenolpyruvate carboxylase, PS00017 ATP/GTP-binding
                  site motif A (P-loop), and PS00393 Phosphoenolpyruvate
                  carboxylase active site 2"
                  /codon_start=1
                  /transl_table=11
                  /product="phosphoenolpyruvate carboxylase"
                  /protein_id="CAB83675.1"
                  /db_xref="GI:7379127"

  Query Match            98.6%;  Score 3160.6;  DB 1;  Length 349061;
  Best Local Similarity  99.3%;  Pred. No. 0;
  Matches 3185;  Conservative    0;  Mismatches   19;  Indels    3;  Gaps    1;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126320 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 126261

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
              ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db     126260 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 126201

Qy        118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     126200 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGGGCAACAACAGCGAAATCAGCA 126141

Qy        178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126140 GCAATATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 126081

Qy        238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126080 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 126021

Qy        298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126020 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 125961

Qy        358 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGC 417
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     125960 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC 125901

Qy        418 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125900 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 125841

Qy        478 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     125840 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGACGGT 125781

Qy        538 AAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 597
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125780 AAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 125721

Qy        598 GATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGG 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125720 GATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGG 125661

Qy        658 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125660 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 125601

Qy        718 AATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTC 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125600 AATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTC 125541

Qy        778 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125540 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCA 125481

Qy        838 GGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125480 GGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 125421

Qy        898 GCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125420 GCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 125361

Qy        958 GGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAAT 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125360 GGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAAT 125301

Qy       1018 GACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAA 1077
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     125300 GACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGACGCTCAA 125241

Qy       1078 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAAC 1137
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     125240 AAAGGCATTATCACAGTTGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAAC 125181

Qy       1138 CATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTC 1197
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     125180 CATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTC 125121

Qy       1198 ACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGC 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125120 ACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGC 125061

Qy       1258 ACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACG 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125060 ACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACG 125001

Qy       1318 CTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125000 CTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGA 124941

Qy       1378 CTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACC 1437
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124940 CTGCTGGATACGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACC 124881

Qy       1438 GCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACG 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124880 GCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACG 124821

Qy       1498 GGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACG 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124820 GGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACG 124761

Qy       1558 GGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATG 1617
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124760 GGCAAAACCGTTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATG 124701

Qy       1618 CGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAAC 1677
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124700 CACGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAAC 124641

Qy       1678 AGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCAC 1737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124640 AGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCAC 124581

Qy       1738 ATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTG 1797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124580 ATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTG 124521

Qy       1798 CTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAA 1857
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||| || 
Db     124520 CTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCAGGCGGTAAG 124461

Qy       1858 GGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATC 1917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124460 GGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATC 124401

Qy       1918 GGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTC 1977
```

```
               ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  124400 GGGCGGGATTATTCTTCTTCACAAACATCGAAACCGACGGCGGTCTGCTGGCTTCCCTC 124341

Qy    1978 GACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGC 2037
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  124340 GACAGCGTCGAAAAAACAGCGGGTAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGC 124281

Qy    2038 GGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACAC 2097
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124280 GGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACAC 124221

Qy    2098 GCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAA 2157
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124220 GCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAA 124161

Qy    2158 TCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGC 2217
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124160 TCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGC 124101

Qy    2218 ATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCC 2277
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124100 ATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCC 124041

Qy    2278 GACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 2337
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  124040 GACGGTGTACGCATCTTCAACAATCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCC 123981

Qy    2338 CATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCT 2397
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  123980 CATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCT 123921

Qy    2398 ACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGT 2457
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123920 ACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGT 123861

Qy    2458 GTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAA 2517
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123860 GTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAA 123801

Qy    2518 AATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCA 2577
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123800 AATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCA 123741

Qy    2578 AATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATC 2637
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123740 AATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATC 123681

Qy    2638 GGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACC 2697
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123680 GGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACC 123621

Qy    2698 GGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTG 2757
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123620 GGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTG 123561

Qy    2758 GGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGC 2817
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123560 GGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGC 123501

Qy    2818 TACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGC 2877
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123500 TACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGC 123441

Qy    2878 AACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTG 2937
           |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123440 AACAGCATCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTG 123381

Qy    2938 AGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGAC 2997
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123380 AGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGAC 123321

Qy    2998 TACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGC 3057
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123320 TACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGC 123261

Qy    3058 AATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGC 3117
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123260 AATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGC 123201

Qy    3118 TGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA 3177
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123200 TGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA 123141

Qy    3178 CGAGTCGGCGTAGGCTACCGGTTCTGA 3204
           |||||||||||||||||||||||||||
Db  123140 CGAGTCGGCGTAGGCTACCGGTTCTGA 123114


RESULT 4
AE002545
LOCUS       AE002545               10591 bp    DNA     linear   BCT 25-MAY-2000
DEFINITION  Neisseria meningitidis serogroup B strain MC58 section 187 of 206
            of the complete genome.
ACCESSION   AE002545 AE002098
VERSION     AE002545.1  GI:7227229
KEYWORDS    .
SOURCE      Neisseria meningitidis MC58.
  ORGANISM  Neisseria meningitidis MC58
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 10591)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Complete genome sequence of Neisseria meningitidis serogroup B
            strain MC58
  JOURNAL   Science 287 (5459), 1809-1815 (2000)
  MEDLINE   20175755
   PUBMED   10710307
REFERENCE   2  (bases 1 to 10591)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..10591
                     /organism="Neisseria meningitidis MC58"
                     /strain="MC58"
                     /db_xref="taxon:122586"
                     /note="serogroup: B"
     gene            548..3796
                     /gene="NMB1969"
     CDS             548..3796
                     /gene="NMB1969"
                     /note="similar to SP:P31631 GB:U07788 GB:M62363 PID:507431
                     percent identity: 48.64; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="serotype-1-specific antigen, putative"
                     /protein_id="AAF42298.1"
                     /db_xref="GI:7227230"
```

```
                /translation="MRTTPTFPTKTFKPTAMALAVATTLSACLGGGGGGTSAPDFNAG
                GTGIGSNSRATTAKSAAVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPPNLH
                TGDFPNPNDAYKNLINLKPAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYN
                ENYKNYTAYMRKEAPEDGGGKDIEASFDDEAVIETEAKPTDIRHVKEIGHIDLVSHII
                GGRSVDGRPAGGIAPDATLHIMNTNDETKNEMMVAAIRNAWVKLGERGVRIVNNSFGT
                TSRAGTADLFQIANSEEQYRQALLDYSGGDKTDEGIRLMQQSDYGNLSYHIRNKNMLF
                IFSTGNDAQAQPNTYALLPFYEKDAQKGIITVAGVDRSGEKFKREMYGEPGTEPLEYG
                SNHCGITAMWCLSAPYEASVRFTRTNPIQIAGTSFSAPIVTGTAALLLQKYPWMSNDN
                LRTTLLTTAQDIGAVGVDSKFGWGLLDAGKAMNGPASFPFGDFTADTKGTSDIAYSFR
                NDISGTGGLIKKGGSQLQLHGNNTYTGKTIIEGGSLVLYGNNKSDMRVETKGALIYNG
                AASGGSLNSDGIVYLADTDQSGANETVHIKGSLQLDGKGTLYTRLGKLLKVDGTAIIG
                GKLYMSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAG
                SEGDTLSYYVRRGNAARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPE
                TVETAAADRTDMPGIRPYGATFRAAAAVQHANAADGVRIFNSLAATVYADSTAAHADM
                QGRRLKAVSDGLDHNGTGLRVIAQTQQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENT
                TAAATLGMGRSTWSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRST
                GADEHAEGSVNGTLMQLGALGGVNVPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGW
                SGNSLTEGTLVGLAGLKLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGK
                TGARNMPHTRLVAGLGADVEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRF"
gene            4828. .6609
                /gene="NMB1970"
CDS             4828. .6609
                /gene="NMB1970"
                /note="similar to GB:AE000511 PID:2313385 percent
                identity: 53.22; identified by sequence similarity;
                putative"
                /codon_start=1
                /transl_table=11
                /product="para-aminobenzoate synthetase component
                I/4-amino-4-deoxychorismate lyase, putative"
                /protein_id="AAF42299.1"
                /db_xref="GI:7227231"
                /translation="MPYFALFDDAVSGRAKRYQNHVESRFFRPEELDALDGALQSGWQ
                KGLHSVLFADYGFGLPLTGVESERGGNLALHWFANCADIDAESWLARHSDGLPAGIST
                PQPSVSETDYLDRIRQIHEAIRRGDTYQINYTTRLHLQAYGNPVSLYRRLRQPVPYAV
                LSHLPDAEGQSAWTLCFSPELFLKIGSDGTISTEPMKGTAPILGDGQDERRAAELQAD
                PKNRAENVMIVDLLRNDLGKIAQTGTVCVPEPFKVSRFGSVWQMTSTIQAQALPHTSF
                ADILRAAFPCGSITGAPKKMSMQIIESLEAEARGLYTGSIGYLNPCSGGLGFEGTFNV
                VIRTLSLTPLSDGIYQGVYGVGSGIVIDSDPAAEYRECGWKARFLNELRPDFGIFETL
                RAENGRCTLLDRHLCRLKTSAQALNLPLPDGCENQIKQYIADLPDGAFRVKALLASDG
                ISLSRAVLNRLTDKQRVIISPAVLPAQNYLRRFKTTCRALFDQAWQTAETQGAFDSLF
                FNSDGILLEGGRSNVFIKHRGQWLTPSLDLDILNGIMRQAVLDEPQKYLQTNQVIETH
                ITQKTLQEAEEIRLSNALRGVFAAALA"
gene            complement(6668. .8185)
                /gene="NMB1971"
CDS             complement(6668. .8185)
                /gene="NMB1971"
                /note="conserved hypothetical protein; identified by
                Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAF42300.1"
                /db_xref="GI:7227232"
                /translation="MLYFRYGFLVVWCAAGVSAAYGADAPAILDDKALLQVQRSVSDK
                WAESDWKVENDAPRVVDGDFLLAHPKMLEHSLRDALNGNQADLIASLADLYAKLPDYD
                AVLYGRARALLAKLAGRPAEAVARYRELHGENAADERILLDLAAAEFDDFRLKSAERH
                FAEAAKLDLPAPVLENVGRFRKKTEGLTGWRFSGGISPAVNRNANNAAPQYCRQNGGR
                QICSVSRAERAAGLNYEIEAEKLTPLADNHYLLFRSNIGGTSYYFSKKSAYDDGFGRA
                YLGWQYKNARQTAGILPFYQVQLSGSDGFDAKTKRVNNRRLPPYMLAHGVGVQLSHTY
                RPNPGWQFSVALEHYRQRYREQDRAEYNNGRQDGFYVSSAKRLGESATVFGGWQFVRF
                VPKRETVGGAVNNAAYRRNGVYAGWAQEWRQLGGLNSRVSASYARRNYKGIAAFSTEA
                QRNREWNVSLALSHDKLSYKGIVPALNYRFGRTESNVPYAKRRNSEVFVSADWRF"
gene            complement(8881. .10515)
                /gene="NMB1972"
CDS             complement(8881. .10515)
                /gene="NMB1972"
                /note="similar to SP:P29842 GB:Z23008 GB:Z23009 PID:438185
                PID:438187 percent identity: 99.26; identified by sequence
                similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="chaperonin, 60 kDa"
```

```
                /protein_id="AAF42301.1"
                /db_xref="GI:7227233"
                /translation="MAAKDVQFGNEVRQKMVNGVNILANAVRVTLGPKGRNVVVDRAF
                GGPHITKDGVTVAKEIELKDKFENMGAQMVKEVASKTNDVAGDGTTTATVLAQSIVAE
                GMKYVTAGMNPTDLKRGIDKAVAALVDELKNIAKPCDTSKEIAQVGSISANSDEQVGA
                IIAEAMEKVGKEGVITVEDGKSLENELDVVEGMQFDRGYLSPYFINDAEKQIAALDNP
                FVLLFDKKISNIRDLLPVLEQVAKASRPLLIIAEDVEGEALATLVVNNIRGILKTVAV
                KAPGFGDRRKAMLQDIAILTGGVVISEEVGLSLEKATLDDLGQAKRIEIGKENTTIID
                GFGDAAQIEARVAEIRQQIETATSDYDKEKLQERVAKLAGGVAVIKVGAATEVEMKEK
                KDRVEDALHATRAAVEEGVVAGGGVALLRARAALENLHTGNADQDAGVQIVLRAVESP
                LRQIVANAGGEPSVVVNKVLEGKGNYGYNAGSGEYGDMIEMGVLDPAKVTRSALQHAA
                SIAGLMLTTDCMIAEIPEDKPAVPDMGGMGGMGGMM"
BASE COUNT     2623 a   3083 c   2609 g   2276 t
ORIGIN

  Query Match            95.6%;  Score 3062.2;  DB 1;  Length 10591;
  Best Local Similarity  97.1%;  Pred. No. 0;
  Matches 3156;  Conservative    0;  Mismatches   48;  Indels   45;  Gaps    2;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        548 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGACTGCCATGGCGTTAGCT 607

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
              ||||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        608 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 667

Qy        118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        668 TTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 727

Qy        178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 787

Qy        238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 847

Qy        298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 357
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        848 CTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 907

Qy        358 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGC 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        908 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC 967

Qy        418 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 1027

Qy        478 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 1087

Qy        538 AAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 597
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1088 AAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 1147

Qy        598 GATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGG 657
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       1148 GATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGG 1207

Qy        658 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 1267

Qy        718 AATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTC 777
              |||||| |||||| ||||||||||||||| ||| | ||||||||||||||||||||||||
Db       1268 AATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTC 1327

Qy        778 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1328 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCA 1387

Qy    838 GGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 897
Db   1388 GGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 1447

Qy    898 GCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 957
Db   1448 GACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 1507

Qy    958 GGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAAT 1017
Db   1508 GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAAT 1567

Qy   1018 GACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAA 1077
Db   1568 GACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAA 1627

Qy   1078 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA----------- 1126
Db   1628 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATG 1687

Qy   1127 ------------------------------ATGGCTCCAACCATTGCGGAATTACTGCC 1155
Db   1688 TATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCC 1747

Qy   1156 ATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 1215
Db   1748 ATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 1807

Qy   1216 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 1275
Db   1808 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 1867

Qy   1276 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAG 1335
Db   1868 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAG 1927

Qy   1336 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 1395
Db   1928 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 1987

Qy   1396 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 1455
Db   1988 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 2047

Qy   1456 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 1515
Db   2048 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 2107

Qy   1516 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 1575
Db   2108 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 2167

Qy   1576 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 1635
Db   2168 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 2227

Qy   1636 GCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTAT 1695
Db   2228 GCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT 2287

Qy   1696 CTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAG 1755
Db   2288 CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAG 2347

Qy   1756 CTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 1815
Db   2348 CTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 2407

Qy   1816 GCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAAC 1875
Db   2408 GCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAAC 2467

Qy   1876 CGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTC 1935
Db   2468 AGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTC 2527

Qy   1936 TTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 1995
Db   2528 TTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 2587

Qy   1996 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 2055
Db   2588 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 2647

Qy   2056 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 2115
Db   2648 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 2707

Qy   2116 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 2175
Db   2708 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 2767

Qy   2176 ACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 2235
Db   2768 ACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 2827

Qy   2236 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 2295
Db   2828 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 2887

Qy   2296 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 2355
Db   2888 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 2947

Qy   2356 CGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATC 2415
Db   2948 CGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATC 3007

Qy   2416 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 2475
Db   3008 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 3067

Qy   2476 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 2535
Db   3068 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 3127

Qy   2536 ACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 2595
Db   3128 ACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 3187

Qy   2596 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 2655
Db   3188 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 3247

Qy   2656 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 2715
Db   3248 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 3307

Qy   2716 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 2775
Db   3308 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 3367

Qy   2776 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 2835
Db   3368 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 3427

Qy   2836 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 2895
Db   3428 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 3487

Qy   2896 ACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 2955
Db   3488 ACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 3547
```

```
Qy    2956 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 3015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3548 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 3607

Qy    3016 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGC 3075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    3608 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGT 3667

Qy    3076 CTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 3135
           ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    3668 CTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 3727

Qy    3136 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 3195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3728 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 3787

Qy    3196 CGGTTCTGA 3204
           |||||||||
Db    3788 CGGTTCTGA 3796


RESULT 5
AX044034
LOCUS       AX044034               349980 bp    DNA     linear   PAT 24-NOV-2000
DEFINITION  Sequence 113 from Patent WO0066791.
ACCESSION   AX044034
VERSION     AX044034.1  GI:11342918
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 349980)
  AUTHORS   Pizza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
            Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
            Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
  TITLE     Neisseria genomic sequences and methods of their use
  JOURNAL   Patent: WO 0066791-A 113 09-NOV-2000;
            CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
                     /note="sequence too long, cut in 8 pieces.~seq 1: 1 to
                     349980 349980 bases~seq 108: 300001 to 649980 349980
                     bases~seq 109: 600001 to 949980 349980 bases~seq 110:
                     900001 to 1249980 349980 bases~seq 111: 1200001 to 1549980
                     349980 bases~seq 112: 1500001 to 1849980 349980 bases~seq
                     113: 1800001 to 2149980 349980 bases~seq 114: 2100001 to
                     2272325 172325 bases"
BASE COUNT     86473 a  95646 c  85908 g  81953 t
ORIGIN

  Query Match             95.6%;  Score 3062.2;  DB 6;  Length 349980;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3156;  Conservative    0;  Mismatches   48;  Indels   45;  Gaps    2;

Qy       1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  261412 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGACTGCCATGGCGTTAGCT 261471

Qy      61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
           ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db  261472 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 261531

Qy     118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
           |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  261532 TTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 261591

Qy     178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261592 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 261651

Qy     238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261652 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAAT 261711

Qy     298 CTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 357
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  261712 CTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAA 261771

Qy     358 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGC 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  261772 CCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC 261831

Qy     418 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261832 GAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAAC 261891

Qy     478 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261892 GAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGT 261951

Qy     538 AAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 597
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261952 AAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACG 262011

Qy     598 GATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGG 657
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  262012 GATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGG 262071

Qy     658 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 717
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262072 CGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATG 262131

Qy     718 AATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTC 777
           |||||| |||||| ||||||||||||||| |||  |||||||||||||||||||||||||
Db  262132 AATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTC 262191

Qy     778 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGAGGGCA 837
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262192 AAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGAGGGCA 262251

Qy     838 GGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 897
           ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  262252 GGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTC 262311

Qy     898 GCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 957
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262312 GACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC 262371

Qy     958 GGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTCGGCAAGCAAT 1017
           |||||| |||||||||||||||||||||||||||||||||||| ||||| || |||||
Db  262372 GGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTCGACAGGCAAT 262431

Qy    1018 GACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAA 1077
           ||||||||||||||||||||||||    ||||||||||||||||||||||||| ||||||
Db  262432 GACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAA 262491

Qy    1078 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA----------- 1126
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  262492 AAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATG 262551

Qy    1127 -------------------------------ATGGCTCCAACCATTGCGGAATTACTGCC 1155
                                          |||||||||||||||||||||||||||||
Db  262552 TATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCC 262611

Qy    1156 ATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 1215
           ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  262612 ATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATT 262671

Qy    1216 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 1275
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262672 CAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTG 262731

Qy    1276 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAG 1335
```

```
                 |||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db    262732 CAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAG 262791

Qy      1336 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 1395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262792 GACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAG 262851

Qy      1396 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 1455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262852 GCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACA 262911

Qy      1456 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 1515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262912 TCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAA 262971

Qy      1516 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 1575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262972 GGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAA 263031

Qy      1576 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 1635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263032 GGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGT 263091

Qy      1636 GCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTAT 1695
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    263092 GCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT 263151

Qy      1696 CTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAG 1755
             |||||||||||||||| ||||||||||||||||||||| ||||||||||||  |||||||
Db    263152 CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAG 263211

Qy      1756 CTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 1815
             |||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263212 CTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACG 263271

Qy      1816 GCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAAC 1875
             ||||| | ||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    263272 GCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAAC 263331

Qy      1876 CGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGCGGGATTATTCTTTC 1935
              ||||||||| ||||||||||||||||||||||||||||||||||| ||||||||||||
Db    263332 AGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGCAGGATTATTCTTTC 263391

Qy      1936 TTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 1995
             ||||||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db    263392 TTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACA 263451

Qy      1996 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263452 GCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACT 263511

Qy      2056 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263512 GCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGC 263571

Qy      2116 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263572 AGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAG 263631

Qy      2176 ACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 2235
             ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    263632 ACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCA 263691

Qy      2236 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 2295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263692 ACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTC 263751

Qy      2296 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 2355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263752 AACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGA 263811

Qy      2356 CGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATC 2415
             ||||| ||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db    263812 CGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATC 263871

Qy      2416 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 2475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263872 GCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC 263931

Qy      2476 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 2535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263932 GGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCC 263991

Qy      2536 ACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 2595
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    263992 ACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGC 264051

Qy      2596 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 2655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264052 ATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTG 264111

Qy      2656 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 2715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264112 TTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCG 264171

Qy      2716 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 2775
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264172 GAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCG 264231

Qy      2776 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 2835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264232 TTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAG 264291

Qy      2836 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 2895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264292 GATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGC 264351

Qy      2896 ACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 2955
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264352 ACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTG 264411

Qy      2956 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 3015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264412 TTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGC 264471

Qy      3016 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGC 3075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    264472 TTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGT 264531

Qy      3076 CTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 3135
             ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    264532 CTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGT 264591

Qy      3136 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 3195
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264592 TACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTAC 264651

Qy      3196 CGGTTCTGA 3204
             |||||||||
Db    264652 CGGTTCTGA 264660
```

```
RESULT 6
NME277537
LOCUS       NME277537              3254 bp    DNA     linear   BCT 24-JUL-2002
DEFINITION  Neisseria meningitidis ausP gene for autotransporter serine
            protease.
ACCESSION   AJ277537
VERSION     AJ277537.2  GI:13539237
KEYWORDS    ausP gene; autotransporter serine protease.
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1
  AUTHORS   Turner,D.P., Wooldridge,K.G. and Ala'Aldeen,D.A.
```

```
TITLE       Autotransported Serine Protease A of Neisseria meningitidis: an
            Immunogenic, Surface-Exposed Outer Membrane, and Secreted Protein
  JOURNAL   Infect. Immun. 70 (8), 4447-4461 (2002)
  MEDLINE   22112893
REFERENCE   2
  AUTHORS   Ala'Aldeen,D.A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Ala'Aldeen D.A.A., Microbiology, University
            of Nottingham, Meningococcal Research Group, Division of
            Microbiology, University Hospital, Nottingham NG7 2UH, UNITED
            KINGDOM
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 3254)
  AUTHORS   Ala'Aldeen,D.A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2001) Ala'Aldeen D.A.A., Microbiology, University
            of Nottingham, Meningococcal Research Group, Division of
            Microbiology, University Hospital, Nottingham NG7 2UH, UNITED
            KINGDOM
COMMENT     On Apr 3, 2001 this sequence version replaced gi:7649689.
FEATURES             Location/Qualifiers
     source          1..3254
                     /organism="Neisseria meningitidis"
                     /strain="B:15:P1.16"
                     /isolate="SD"
                     /db_xref="taxon:487"
                     /country="United Kingdom:England"
     gene            1..3254
                     /gene="ausP"
                     /pseudo
     CDS             1..3254
                     /gene="ausP"
                     /function="unknown"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="autotransporter serine protease"
BASE COUNT      845 a    922 c    917 g    570 t
ORIGIN

  Query Match             95.0%;  Score 3042.4;  DB 1;  Length 3254;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 3153;  Conservative    0;  Mismatches   51;  Indels   50;  Gaps    3;

Qy          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGGCTGCCATGGCGTTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 ATGCGAACGACCCCAACCTTCCCTACAAAAACTTTCAAACCGACTGCCATGGCGTTAGCT 60

Qy         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGC---GGCGGCACTTCTGCGCCCGAC 117
              ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db         61 GTTGCAACAACACTTTCTGCCTGCTTAGGCGGCGGCGGAGGCGGCACTTCTGCGCCCGAC 120

Qy        118 TTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 177
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA 180

Qy        178 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCC 240

Qy        238 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATG-----CCCCCCCCC 292
              ||||||||||||||||||||||||||||||||||||||||||||||     |||||||||
Db        241 GGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCCCCC 300

Qy        293 CGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACC 352
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        301 CGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTGATCAACC 360

Qy        353 TCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATA 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        361 TCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACA 420

Qy        413 CAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCT 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCT 480

Qy        473 ATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAG 532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAG 540

Qy        533 GCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGC 592
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGC 600

Qy        593 CGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTG 652
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        601 CGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTG 660

Qy        653 GCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACA 712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACA 720

Qy        713 TAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCAT 772
              ||||||||||| ||||||||||||||||||||||| |||  |||||||||||||||||||
Db        721 TAATGAATACGAATGATGGAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCAT 780

Qy        773 GGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGA 832
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTGGAACAACATCGA 840

Qy        833 GGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGT 892
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGT 900

Qy        893 TGCTCGCCTATTCCGGCGGTGATAAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCG 952
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCTCGACTATTCCGGCGGTGATAAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCG 960

Qy        953 ATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAA 1012
              ||||||||||| ||||||||||||||||||||||||||||||||||||| ||||| ||
Db        961 ATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAG 1020

Qy       1013 GCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTATGAAAAAGATG 1072
              |||||||||||||||||||||||||||       |||||||||||||||||||||||| |
Db       1021 GCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTATGAAAAAGACG 1080

Qy       1073 CTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA------ 1126
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1081 CTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGCGGAGAAAAGTTCAAACGGG 1140

Qy       1127 ------------------------------------ATGGCTCCAACCATTGCGGAATTA 1150
                                                  ||||||||||||||||||||||||
Db       1141 AAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTA 1200

Qy       1151 CTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACC 1210
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACC 1260

Qy       1211 CGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGC 1320

Qy       1271 TGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGG 1330
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||
Db       1321 TGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGG 1380

Qy       1331 CTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGG 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGG 1440

Qy       1391 GTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAG 1450
              |||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTAAGGCCTTGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAG 1500

Qy       1451 GTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCA 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1501 GTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCA 1560

Qy   1511 AAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTA 1570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTA 1620

Qy   1571 TCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCA 1630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCA 1680

Qy   1631 AAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTG 1690
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1681 AAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTG 1740

Qy   1691 TCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATC 1750
          ||||||||||||||||||||| ||||||||||||||||||||| ||||||||||||  ||
Db   1741 TCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTC 1800

Qy   1751 TGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACG 1810
          ||||||||| |||| |||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACG 1860

Qy   1811 GTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATC 1870
          |||||||||| | ||||||||||||||||||||||||||||||||| |||||||||||||
Db   1861 GTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATC 1920

Qy   1871 TCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCGGGATTATT 1930
           |||| ||||||||| ||||||||||||||||||||||||||||||||||| ||||||||
Db   1921 ACAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATT 1980

Qy   1931 CTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAA 1990
          |||||||||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db   1981 CTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAA 2040

Qy   1991 AAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCAC 2050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCAC 2100

Qy   2051 GGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGG 2110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGG 2160

Qy   2111 GCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACAC 2170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCCGCAACAC 2220

Qy   2171 CCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACG 2230
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   2221 CCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACG 2280

Qy   2231 GCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCA 2290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCA 2340

Qy   2291 TCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGC 2350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGC 2400

Qy   2351 AGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCG 2410
          |||||||||| ||||||||||||||||||||||||||||||||||  |||||||||||||
Db   2401 AGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCG 2460

Qy   2411 TCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAA 2470
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   2461 TCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTTGAGGGCAAAA 2520

Qy   2471 TGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAG 2530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAG 2580

Qy   2531 CCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCG 2590
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2581 CCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCG 2640

Qy   2591 ACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAG 2650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 ACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTCAAAG 2700

Qy   2651 GCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAAC 2710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAAC 2760

Qy   2711 ATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACG 2770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 ATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACG 2820

Qy   2771 TTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCA 2830
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCA 2880

Qy   2831 AACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTG 2890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTG 2940

Qy   2891 AAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCG 2950
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 AAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCG 3000

Qy   2951 TCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGG 3010
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 TCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGG 3060

Qy   3011 GCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACA 3070
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACA 3120

Qy   3071 CCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGG 3130
          |||| ||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   3121 CCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGG 3180

Qy   3131 CACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAG 3190
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 CACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAG 3240

Qy   3191 GCTACCGGTTCTGA 3204
          ||||||||||||||
Db   3241 GCTACCGGTTCTGA 3254
```

```
RESULT 7
AX236469
LOCUS       AX236469                4218 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 162 from Patent WO0164922.
ACCESSION   AX236469
VERSION     AX236469.1  GI:15796056
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4218)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 162 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4218
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961cL-983"
BASE COUNT     1196 a   1149 c   1130 g    743 t
ORIGIN

  Query Match            93.2%;  Score 2986.2;  DB 6;  Length 4218;
  Best Local Similarity  97.1%;  Pred. No. 0;
  Matches 3067;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;
```

```
Qy     90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
Db   1056 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1115

Qy    150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
Db   1116 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1175

Qy    210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
Db   1176 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1235

Qy    270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
Db   1236 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1295

Qy    330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
Db   1296 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1355

Qy    390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
Db   1356 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1415

Qy    450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
Db   1416 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1475

Qy    510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
Db   1476 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1535

Qy    570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
Db   1536 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1595

Qy    630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
Db   1596 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1655

Qy    690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
Db   1656 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1715

Qy    750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
Db   1716 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1775

Qy    810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
Db   1776 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1835

Qy    870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
Db   1836 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1895

Qy    930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
Db   1896 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1955

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
Db   1956 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 2015

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
Db   2016 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2075

Qy   1110 CAGTGGAGAAAAGTTCA-----------------------------------------A 1127
Db   2076 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2135

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
Db   2136 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2195

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
Db   2196 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2255

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
Db   2256 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2315

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
Db   2316 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2375

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
Db   2376 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2435

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
Db   2436 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2495

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
Db   2496 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2555

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
Db   2556 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2615

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
Db   2616 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2675

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
Db   2676 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2735

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
Db   2736 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2795

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
Db   2796 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2855

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
Db   2856 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2915

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
Db   2916 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2975

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
Db   2976 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 3035

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
Db   3036 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3095

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
Db   3096 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3155

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
Db   3156 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3215

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3216 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3275

Qy       2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3276 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3335

Qy       2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       3336 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3395

Qy       2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3396 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3455

Qy       2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3456 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3515

Qy       2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db       3516 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3575

Qy       2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3576 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3635

Qy       2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3636 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3695

Qy       2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3696 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3755

Qy       2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3756 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3815

Qy       2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3816 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3875

Qy       2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       3876 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3935

Qy       2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3936 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3995

Qy       2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3996 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 4055

Qy       3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
              ||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
Db       4056 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4115

Qy       3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4116 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4175

Qy       3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
              |||||||||||||||||||||||||||||||||||||
Db       4176 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212

RESULT 8
AX239755
LOCUS       AX239755               4218 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 57 from Patent WO0164920.
ACCESSION   AX239755
VERSION     AX239755.1  GI:15797388
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4218)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 57 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4218
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961cL-983"
BASE COUNT     1196 a   1149 c   1130 g    743 t
ORIGIN

  Query Match             93.2%;  Score 2986.2;  DB 6;  Length 4218;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3067;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy         90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
              |||||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1056 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1115

Qy        150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1116 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1175

Qy        210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1176 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1235

Qy        270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db       1236 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1295

Qy        330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1296 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1355

Qy        390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1356 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1415

Qy        450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1416 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1475

Qy        510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1476 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1535

Qy        570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1536 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1595

Qy        630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1596 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1655

Qy        690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
              |||||||||||||||||||||||||||||||||| ||||| |||||||||||||||| ||
Db       1656 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1715

Qy        750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
              |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1716 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1775

Qy        810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
```

```
Db   1776 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1835

Qy    870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
Db   1836 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1895

Qy    930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
Db   1896 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1955

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
Db   1956 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 2015

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
Db   2016 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2075

Qy   1110 CAGTGGAGAAAAGTTCA---------------------------------------A 1127
Db   2076 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2135

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
Db   2136 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2195

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
Db   2196 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2255

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
Db   2256 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2315

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
Db   2316 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2375

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
Db   2376 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2435

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
Db   2436 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2495

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
Db   2496 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2555

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
Db   2556 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2615

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
Db   2616 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2675

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
Db   2676 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2735

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
Db   2736 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2795

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
Db   2796 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2855

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
Db   2856 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2915

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
Db   2916 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2975

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
Db   2976 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 3035

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
Db   3036 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3095

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
Db   3096 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3155

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
Db   3156 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3215

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
Db   3216 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3275

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
Db   3276 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3335

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
Db   3336 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3395

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
Db   3396 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3455

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
Db   3456 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3515

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
Db   3516 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3575

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
Db   3576 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3635

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
Db   3636 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3695

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
Db   3696 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3755

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
Db   3756 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3815

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
Db   3816 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3875

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
Db   3876 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3935

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
Db   3936 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3995
```

```
Qy     2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
Db     3996 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 4055

Qy     3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
Db     4056 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4115

Qy     3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
Db     4116 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4175

Qy     3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
Db     4176 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTCTGA 4212


RESULT 9
AX236457
LOCUS       AX236457                4335 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 150 from Patent WO0164922.
ACCESSION   AX236457
VERSION     AX236457.1  GI:15796050
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4335)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 150 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961-983"
BASE COUNT     1217 a   1189 c   1165 g    764 t
ORIGIN

  Query Match             93.1%;  Score 2983.8;  DB 6;  Length 4335;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches   57;  Indels   42;  Gaps    1;

Qy       75 TTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCAC 134
Db     1140 TTACGAGTGGGGATCCGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTAC 1199

Qy      135 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 194
Db     1200 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 1259

Qy      195 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 254
Db     1260 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 1319

Qy      255 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 314
Db     1320 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 1379

Qy      315 TACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 374
Db     1380 TCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 1439

Qy      375 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCAT 434
Db     1440 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCAT 1499

Qy      435 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 494
Db     1500 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 1559

Qy      495 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTC 554
Db     1560 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTC 1619

Qy      555 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 614
Db     1620 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 1679

Qy      615 AGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 674
Db     1680 AGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 1739

Qy      675 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAAC 734
Db     1740 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAAC 1799

Qy      735 CAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 794
Db     1800 CAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 1859

Qy      795 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTT 854
Db     1860 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTT 1919

Qy      855 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGA 914
Db     1920 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGA 1979

Qy      915 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCA 974
Db     1980 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCA 2039

Qy      975 CATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCC 1034
Db     2040 CATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCC 2099

Qy     1035 CAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGT 1094
Db     2100 CAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGT 2159

Qy     1095 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA--------------------------- 1126
Db     2160 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTAC 2219

Qy     1127 -------------ATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGC 1172
Db     2220 AGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGC 2279

Qy     1173 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 1232
Db     2280 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 2339

Qy     1233 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 1292
Db     2340 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 2399

Qy     1293 GAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGG 1352
Db     2400 GAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGG 2459

Qy     1353 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 1412
Db     2460 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 2519

Qy     1413 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 1472
Db     2520 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 2579

Qy     1473 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 1532
Db     2580 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 2639

Qy     1533 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 1592
```

```
Db   2640 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 2699

Qy   1593 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 1652
Db   2700 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 2759

Qy   1653 GGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCG 1712
Db   2760 GGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCA 2819

Qy   1713 ATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTAC 1772
Db   2820 ATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTAC 2879

Qy   1773 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAA 1832
Db   2880 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAA 2939

Qy   1833 GCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGT 1892
Db   2940 GCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGT 2999

Qy   1893 TCCCTTCCTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAAC 1952
Db   3000 TCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAAC 3059

Qy   1953 CGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 2012
Db   3060 CGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 3119

Qy   2013 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 2072
Db   3120 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 3179

Qy   2073 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 2132
Db   3180 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 3239

Qy   2133 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 2192
Db   3240 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 3299

Qy   2193 CGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 2252
Db   3300 AGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 3359

Qy   2253 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 2312
Db   3360 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 3419

Qy   2313 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGT 2372
Db   3420 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGT 3479

Qy   2373 ATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 2432
Db   3480 ATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 3539

Qy   2433 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 2492
Db   3540 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 3599

Qy   2493 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACA 2552
Db   3600 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACG 3659

Qy   2553 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 2612
Db   3660 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 3719

Qy   2613 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 2672
Db   3720 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 3779

Qy   2673 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 2732
Db   3780 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 3839

Qy   2733 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 2792
Db   3840 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 3899

Qy   2793 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 2852
Db   3900 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 3959

Qy   2853 AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGC 2912
Db   3960 AGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGC 4019

Qy   2913 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 2972
Db   4020 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 4079

Qy   2973 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 3032
Db   4080 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 4139

Qy   3033 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGG 3092
Db   4140 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGG 4199

Qy   3093 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 3152
Db   4200 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 4259

Qy   3153 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
Db   4260 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308
```

```
RESULT 10
AX239743
LOCUS       AX239743                4335 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 45 from Patent WO0164920.
ACCESSION   AX239743
VERSION     AX239743.1  GI:15797382
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4335)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 45 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961-983"
BASE COUNT     1217 a   1189 c    1165 g     764 t
ORIGIN

  Query Match             93.1%;  Score 2983.8;  DB 6;  Length 4335;
  Best Local Similarity   96.9%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches   57;  Indels   42;  Gaps    1;

Qy     75 TTCTGCCTGCTTAGGCGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCAC 134
Db   1140 TTACGAGTGGGGATCCGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTAC 1199

Qy    135 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 194
```

```
Db    1200 CGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGG 1259

Qy     195 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 254
Db    1260 TATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGC 1319

Qy     255 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 314
Db    1320 GGTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTT 1379

Qy     315 TACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 374
Db    1380 TCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGG 1439

Qy     375 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCAT 434
Db    1440 CTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCAT 1499

Qy     435 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 494
Db    1500 ATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTA 1559

Qy     495 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTC 554
Db    1560 TACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTC 1619

Qy     555 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 614
Db    1620 TTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAA 1679

Qy     615 AGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 674
Db    1680 AGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAG 1739

Qy     675 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAAC 734
Db    1740 ACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAAC 1799

Qy     735 CAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 794
Db    1800 CAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGG 1859

Qy     795 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTT 854
Db    1860 CGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTT 1919

Qy     855 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGA 914
Db    1920 CCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGA 1979

Qy     915 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCA 974
Db    1980 TAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCA 2039

Qy     975 CATCCGTAATAAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCC 1034
Db    2040 CATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCC 2099

Qy    1035 CAACACACTGACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGT 1094
Db    2100 CAACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGT 2159

Qy    1095 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCA---------------------------- 1126
Db    2160 CGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTAC 2219

Qy    1127 --------------ATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGC 1172
Db    2220 AGAACCGCTTGAGTATGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGC 2279

Qy    1173 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 1232
Db    2280 ACCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATC 2339

Qy    1233 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 1292
Db    2340 CTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGAT 2399

Qy    1293 GAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGG 1352
Db    2400 GAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGG 2459

Qy    1353 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 1412
Db    2460 CGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGC 2519

Qy    1413 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 1472
Db    2520 GTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTC 2579

Qy    1473 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 1532
Db    2580 CTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCA 2639

Qy    1533 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 1592
Db    2640 ACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTT 2699

Qy    1593 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 1652
Db    2700 GTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGG 2759

Qy    1653 GGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCG 1712
Db    2760 GGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCA 2819

Qy    1713 ATCCGGCGCAAACGAAACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTAC 1772
Db    2820 ATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTAC 2879

Qy    1773 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAA 1832
Db    2880 GCTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAA 2939

Qy    1833 GCTGTACATGTCGGCACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGT 1892
Db    2940 GCTGTACATGTCGGCACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGT 2999

Qy    1893 TCCCTTCCTGAGTGCCGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAAC 1952
Db    3000 TCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAAC 3059

Qy    1953 CGACGGTGGTCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 2012
Db    3060 CGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGA 3119

Qy    2013 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 2072
Db    3120 CACGCTGTCCTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACA 3179

Qy    2073 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 2132
Db    3180 TTCCGCGCCCGCCGGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCT 3239

Qy    2133 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 2192
Db    3240 GATGGTCGAACTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGC 3299

Qy    2193 CGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 2252
Db    3300 AGCCGACCGCACAGATATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGC 3359

Qy    2253 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 2312
Db    3360 AGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTAC 3419
```

```
Qy     2313 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGACGCCGGCTGAAAGCCGT 2372
            ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     3420 CGTCTATGCCGACAGTACCGCCGCCCATGCCGATATGCAGGACGCCGCCTGAAAGCCGT 3479

Qy     2373 ATCGGACGGGTTGGACCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 2432
            |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db     3480 ATCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGA 3539

Qy     2433 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 2492
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3540 CGGTGGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGT 3599

Qy     2493 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACA 2552
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3600 CGGCATTGCCGCGAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACG 3659

Qy     2553 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 2612
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3660 CAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGG 3719

Qy     2613 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 2672
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3720 CATACGGCACGATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTA 3779

Qy     2673 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 2732
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3780 CAAAAACAGCATCAGCCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGG 3839

Qy     2733 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 2792
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3840 CACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGA 3899

Qy     2793 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 2852
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3900 TTTGACGGTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAA 3959

Qy     2853 AGGCAGTGCTTTGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGC 2912
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     3960 AGGCAGTGCTTTGGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGC 4019

Qy     2913 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 2972
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4020 GGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGT 4079

Qy     2973 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 3032
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4080 GGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGC 4139

Qy     3033 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGG 3092
            |||||||||||||||||||||||||||||||||||||||||| ||||||||||| |||||
Db     4140 AGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGG 4199

Qy     3093 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 3152
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4200 CGCGGATGTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTC 4259

Qy     3153 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db     4260 CAAACAGTACGGCAACCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4308
```

```
RESULT 11
AX236419
LOCUS       AX236419                3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 112 from Patent WO0164922.
ACCESSION   AX236419
VERSION     AX236419.1  GI:15796032
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 112 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG741-983"
BASE COUNT     1046 a   1103 c   1106 g    684 t
ORIGIN

  Query Match             93.1%;  Score 2983.2;  DB 6;  Length 3939;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy       90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
            |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      759 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 818

Qy      150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      819 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 878

Qy      210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      879 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 938

Qy      270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db      939 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 998

Qy      330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      999 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1058

Qy      390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
            ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     1059 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1118

Qy      450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1119 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1178

Qy      510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
            ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     1179 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1238

Qy      570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1239 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1298

Qy      630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
            |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1299 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1358

Qy      690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
            |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||| ||
Db     1359 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1418

Qy      750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
            |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1419 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1478

Qy      810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
            ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     1479 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1538

Qy      870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     1539 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1598

Qy      930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
```

```
Db   1599 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAA 1658

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
Db   1659 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1718

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
Db   1719 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1778

Qy   1110 CAGTGGAGAAAAGTTCA-----------------------------------------A 1127
Db   1779 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 1838

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
Db   1839 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 1898

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
Db   1899 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1958

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
Db   1959 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2018

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
Db   2019 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2078

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
Db   2079 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2138

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
Db   2139 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2198

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
Db   2199 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2258

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
Db   2259 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2318

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
Db   2319 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2378

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
Db   2379 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2438

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
Db   2439 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2498

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
Db   2499 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2558

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
Db   2559 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2618

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
Db   2619 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2678

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
Db   2679 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2738

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
Db   2739 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2798

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
Db   2799 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2858

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
Db   2859 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 2918

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
Db   2919 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2978

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
Db   2979 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3038

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
Db   3039 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3098

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
Db   3099 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3158

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
Db   3159 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3218

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
Db   3219 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3278

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
Db   3279 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3338

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
Db   3339 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3398

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
Db   3399 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3458

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
Db   3459 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3518

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
Db   3519 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3578

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
Db   3579 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3638

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
Db   3639 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3698

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
Db   3699 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3758

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
Db   3759 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 3818
```

```
Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3819 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3878

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          ||||||||||||||||||||||||||||||||||
Db   3879 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912


RESULT 12
AX239729
LOCUS       AX239729                3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 31 from Patent WO0164920.
ACCESSION   AX239729
VERSION     AX239729.1  GI:15797375
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 31 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG741-983"
BASE COUNT     1046 a   1103 c   1106 g    684 t
ORIGIN

  Query Match             93.1%;  Score 2983.2;  DB 6;  Length 3939;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy     90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
          |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    759 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 818

Qy    150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    819 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 878

Qy    210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    879 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 938

Qy    270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    939 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 998

Qy    330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    999 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1058

Qy    390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1059 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1118

Qy    450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1119 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1178

Qy    510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1179 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1238

Qy    570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1239 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1298

Qy    630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1299 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1358

Qy    690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
          |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||| ||
Db   1359 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1418

Qy    750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
          |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1419 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1478

Qy    810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1479 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1538

Qy    870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1539 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1598

Qy    930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1599 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1658

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
          |||||||||||| |||||| || |||||||||||||||||||||||||||||    ||||
Db   1659 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1718

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1719 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1778

Qy   1110 CAGTGGAGAAAAGTTCA------------------------------------------A 1127
          |||||||||||||||||                                          |
Db   1779 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 1838

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1839 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 1898

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1899 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1958

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1959 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2018

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
          |||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||||
Db   2019 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2078

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2079 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2138

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2139 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2198

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2199 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2258

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2259 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2318

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2319 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2378
```

```
Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
           ||||||||||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db   2379 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2438

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
          ||||| |||||||||||| |||||||||||| |||| ||||||||||||||||||||||
Db   2439 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2498

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
          |||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db   2499 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2558

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
          |||||||| ||||||||||||||||| |||||||| ||||||||||||||||||||||
Db   2559 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2618

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
          ||||||||||||| |||||||||||||||||||||||||||||||||||| || |||||
Db   2619 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2678

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2738

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2798

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2858

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   2859 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 2918

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2919 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2978

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3038

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   3039 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3098

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
          |||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3099 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3158

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3218

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   3219 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3278

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3338

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3339 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3398

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3399 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3458

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3459 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3518

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAGGCAGTGCTTTGGG 2867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3519 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAGGCAGTGCTTTGGG 3578

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   3579 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3638

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3639 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3698

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3699 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3758

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
          |||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||||
Db   3759 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 3818

Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3819 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3878

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
          ||||||||||||||||||||||||||||||||||
Db   3879 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3912


RESULT 13
AX236463
LOCUS       AX236463                4170 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 156 from Patent WO0164922.
ACCESSION   AX236463
VERSION     AX236463.1  GI:15796053
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4170)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 156 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4170
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961c-983"
BASE COUNT     1185 a   1138 c   1119 g    728 t
ORIGIN

  Query Match             93.1%;  Score 2983.2;  DB 6;  Length 4170;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy     90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
          ||||||  ||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    990 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1049

Qy    150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1109

Qy    210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1169

Qy    270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
```

```
Db   1170 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1229

Qy    330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
Db   1230 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1289

Qy    390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
Db   1290 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1349

Qy    450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
Db   1350 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1409

Qy    510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
Db   1410 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1469

Qy    570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
Db   1470 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1529

Qy    630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
Db   1530 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1589

Qy    690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
Db   1590 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1649

Qy    750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
Db   1650 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1709

Qy    810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
Db   1710 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1769

Qy    870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
Db   1770 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1829

Qy    930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
Db   1830 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1889

Qy    990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
Db   1890 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1949

Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
Db   1950 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2009

Qy   1110 CAGTGGAGAAAAGTTCA---------------------------------------A 1127
Db   2010 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2069

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
Db   2070 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2129

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
Db   2130 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2189

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
Db   2190 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2249

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
Db   2250 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2309

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
Db   2310 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2369

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
Db   2370 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2429

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
Db   2430 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2489

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
Db   2490 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2549

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
Db   2550 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2609

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
Db   2610 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2669

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
Db   2670 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2729

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
Db   2730 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2789

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
Db   2790 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2849

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
Db   2850 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2909

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
Db   2910 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2969

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
Db   2970 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3029

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
Db   3030 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3089

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
Db   3090 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3149

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
Db   3150 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3209

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
Db   3210 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3269

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
Db   3270 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3329

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
Db   3330 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3389
```

```
Qy       2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3449

Qy       2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACACAGCACATGGAGCGA 2567
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       3450 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGACGCAGCACATGGAGCGA 3509

Qy       2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3569

Qy       2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3570 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3629

Qy       2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3630 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3689

Qy       2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3690 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3749

Qy       2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3750 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3809

Qy       2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       3810 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3869

Qy       2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3870 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3929

Qy       2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3930 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3989

Qy       3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
              |||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db       3990 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4049

Qy       3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4050 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4109

Qy       3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
              ||||||||||||||||||||||||||||||||||
Db       4110 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143
```

```
RESULT 14
AX239749
LOCUS       AX239749              4170 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 51 from Patent WO0164920.
ACCESSION   AX239749
VERSION     AX239749.1  GI:15797385
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4170)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 51 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4170
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="961c-983"
BASE COUNT     1185 a   1138 c   1119 g    728 t
ORIGIN

  Query Match             93.1%;  Score 2983.2;  DB 6;  Length 4170;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 3064;  Conservative    0;  Mismatches   48;  Indels   42;  Gaps    1;

Qy         90 CGGCGGCGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAA 149
              |||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        990 CGGCGGAGGCGGCACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAA 1049

Qy        150 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1050 CAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAAT 1109

Qy        210 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1110 GTGCAAAGACAGAAGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGA 1169

Qy        270 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGA 329
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db       1170 TGCCAAAATCAATGCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGA 1229

Qy        330 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1230 CGCATACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGG 1289

Qy        390 GGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 449
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1290 GGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACT 1349

Qy        450 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1350 GTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCG 1409

Qy        510 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGC 569
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       1410 GAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGC 1469

Qy        570 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1470 CGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACAT 1529

Qy        630 CGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 689
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1530 CGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTAT 1589

Qy        690 TGCGCCCGATGCGACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAAT 749
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||| ||
Db       1590 TGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGAT 1649

Qy        750 GTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 809
              |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1650 GGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAA 1709

Qy        810 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTC 869
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1710 TAACAGTTTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTC 1769

Qy        870 GGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGG 929
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1770 GGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGG 1829

Qy        930 TATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAA 989
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db       1830 TATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAA 1889

Qy        990 CATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCT 1049
              |||||||||||| |||||| || |||||||||||||||||||||||||||||    ||||
Db       1890 CATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCT 1949
```

```
Qy   1050 ATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 1109
Db   1950 ATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCG 2009

Qy   1110 CAGTGGAGAAAAGTTCA------------------------------------------A 1127
Db   2010 CAGTGGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTA 2069

Qy   1128 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAG 1187
Db   2070 TGGCTCCAACCATTGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAG 2129

Qy   1188 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 1247
Db   2130 CGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCAT 2189

Qy   1248 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 1307
Db   2190 CGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCT 2249

Qy   1308 GCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 1367
Db   2250 GCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTT 2309

Qy   1368 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 1427
Db   2310 CGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGG 2369

Qy   1428 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 1487
Db   2370 CGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACAT 2429

Qy   1488 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 1547
Db   2430 TTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAA 2489

Qy   1548 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 1607
Db   2490 CACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAA 2549

Qy   1608 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 1667
Db   2550 ATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGG 2609

Qy   1668 TAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGA 1727
Db   2610 CAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGA 2669

Qy   1728 AACCGTGCACATCAAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTT 1787
Db   2670 AACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTT 2729

Qy   1788 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGC 1847
Db   2730 GGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGC 2789

Qy   1848 ACGCGGCAAAGGGGCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGC 1907
Db   2790 ACGCGGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGC 2849

Qy   1908 CGCCAAAATCGGGCGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCT 1967
Db   2850 CGCCAAAATCGGGCAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCT 2909

Qy   1968 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2027
Db   2910 GGCTTCCCTCGACAGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTA 2969

Qy   2028 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 2087
Db   2970 TGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGG 3029

Qy   2088 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 2147
Db   3030 TCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGA 3089

Qy   2148 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGA 2207
Db   3090 TGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGA 3149

Qy   2208 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 2267
Db   3150 TATGCCGGGCATCCGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGC 3209

Qy   2268 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 2327
Db   3210 GAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAG 3269

Qy   2328 TACCGCCGCCCATGCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGA 2387
Db   3270 TACCGCCGCCCATGCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGA 3329

Qy   2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
Db   3330 CCACAACGGCACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 3389

Qy   2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
Db   3390 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 3449

Qy   2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
Db   3450 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGA 3509

Qy   2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
Db   3510 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 3569

Qy   2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
Db   3570 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 3629

Qy   2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
Db   3630 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 3689

Qy   2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
Db   3690 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 3749

Qy   2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
Db   3750 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 3809

Qy   2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
Db   3810 CTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTC 3869

Qy   2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
Db   3870 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 3929

Qy   2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
Db   3930 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3989

Qy   3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATT 3107
Db   3990 GGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATT 4049

Qy   3108 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 3167
Db   4050 CGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAA 4109

Qy   3168 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 3201
```

```
Db    4110 CCACAGCGGACGAGTCGGCGTAGGCTACCGGTTC 4143


RESULT 15
AX236409
LOCUS       AX236409                3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 102 from Patent WO0164922.
ACCESSION   AX236409
VERSION     AX236409.1  GI:15796027
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 102 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-741"
BASE COUNT     1046 a   1100 c   1107 g    686 t
ORIGIN

  Query Match            92.8%;  Score 2971.8;  DB 6;  Length 3939;
  Best Local Similarity  97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy     103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
           ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy     163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy     223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy     283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy     343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy     403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy     463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy     523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy     583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy     643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy     703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
           ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db     604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy     763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy     823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy     883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
           |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy     943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db     844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy    1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
           |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db     904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy    1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy    1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
           ||||                                          ||||||||||||||
Db    1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy    1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy    1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy    1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy    1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy    1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy    1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy    1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy    1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy    1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
           ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy    1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
           ||||||||||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db    1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy    1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
           ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db    1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743
```

```
Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
Db   2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
Db   2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
Db   3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
Db   3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy   3181 GTCGGCGTAGGCTACCGGTTC 3201
Db   3124 GTCGGCGTAGGCTACCGGTTC 3144


RESULT 16
AX239717
LOCUS       AX239717                3939 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 19 from Patent WO0164920.
ACCESSION   AX239717
VERSION     AX239717.1  GI:15797370
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3939)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 19 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..3939
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-741"
BASE COUNT     1046 a   1100 c   1107 g    686 t
ORIGIN

  Query Match             92.8%;  Score 2971.8;  DB 6;  Length 3939;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy    103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
Db      4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy    163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
Db     64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy    223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
Db    124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy    283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
Db    184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy    343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
Db    244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303
```

```
Qy    403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
Db    304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy    463 GAACACGGCTATAACGAAAATTACAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
Db    364 GAACACGGCTATAACGAAAATTACAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
Db    424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
Db    484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy    643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
Db    544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy    703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
Db    604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy    763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
Db    664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy    823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
Db    724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy    883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
Db    784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy    943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
Db    844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy   1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
Db    904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy   1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
Db    964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy   1123 TTCA-----------------------------------------ATGGCTCCAACCAT 1140
Db   1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
```

```
Db     2464 ACGACAGCAGCCGCCACACTGGGCATGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy     2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy     2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy     2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy     2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy     2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy     2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
            ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy     2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy     3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy     3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
            |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db     3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy     3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy     3181 GTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||
Db     3124 GTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 17
AX236413
LOCUS       AX236413                4179 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 106 from Patent WO0164922.
ACCESSION   AX236413
VERSION     AX236413.1  GI:15796029
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4179)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 106 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1. .4179
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961c"
BASE COUNT     1187 a   1140 c   1121 g    731 t
ORIGIN

  Query Match             92.8%;  Score 2971.8;  DB 6;  Length 4179;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy      103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy      163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy      223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy      283 GCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
            |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db      184 GCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy      343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy      403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy      463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy      523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy      583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy      643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy      703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
            ||||||||||||||||||||| |||||| ||||||||||||||| ||| |  ||||||||||
Db      604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy      763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy      823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy      883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy      943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
            ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy     1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
            |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db      904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy     1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy     1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
            ||||                                          ||||||||||||||
Db     1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083
```

```
Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
          ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
          |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
          ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db   3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy   3181 GTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||
Db   3124 GTCGGCGTAGGCTACCGGTTC 3144

RESULT 18
AX239721
LOCUS       AX239721              4179 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 23 from Patent WO0164920.
```

```
ACCESSION   AX239721
VERSION     AX239721.1  GI:15797372
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4179)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 23 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4179
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961c"
BASE COUNT     1187 a   1140 c   1121 g    731 t
ORIGIN

  Query Match             92.8%;  Score 2971.8;  DB 6;  Length 4179;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy        103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy        163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy        223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy        283 GCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        184 GCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy        343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy        403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy        463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy        523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy        583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy        643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy        703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
              ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db        604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy        763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy        823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy        883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy        943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db        844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy       1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
              |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db        904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy       1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy       1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
              ||||                                          ||||||||||||||
Db       1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy       1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy       1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db       1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy       1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
              ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db       1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy       1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
              |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db       1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy       1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
              |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
```

```
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
          | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy   2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy   2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy   2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy   2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy   3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy   3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
          |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db   3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy   3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy   3181 GTCGGCGTAGGCTACCGGTTC 3201
          |||||||||||||||||||||
Db   3124 GTCGGCGTAGGCTACCGGTTC 3144


RESULT 19
AX236411
LOCUS       AX236411                4344 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 104 from Patent WO0164922.
ACCESSION   AX236411
VERSION     AX236411.1  GI:15796028
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4344)
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 104 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4344
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961"
BASE COUNT     1219 a   1191 c   1167 g    767 t
ORIGIN

  Query Match             92.8%;  Score 2971.8;  DB 6;  Length 4344;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy    103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy    163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy    223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy    283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy    343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy    403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy    463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423
```

```
Qy    523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
Db    424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy    583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
Db    484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy    643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
Db    544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy    703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
Db    604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy    763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
Db    664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy    823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
Db    724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy    883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
Db    784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy    943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
Db    844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy   1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
Db    904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy   1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
Db    964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy   1123 TTCA----------------------------------------ATGGCTCCAACCAT 1140
Db   1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy     2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy     2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy     2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy     2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
            ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy     2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy     3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy     3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
            |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db     3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy     3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy     3181 GTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||
Db     3124 GTCGGCGTAGGCTACCGGTTC 3144
```

```
RESULT 20
AX239719
LOCUS       AX239719                4344 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 21 from Patent WO0164920.
ACCESSION   AX239719
VERSION     AX239719.1  GI:15797371
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4344)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 21 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4344
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-961"
BASE COUNT     1219 a   1191 c   1167 g    767 t
ORIGIN

  Query Match             92.8%;  Score 2971.8;  DB 6;  Length 4344;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy      103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy      163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy      223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy      283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
            ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy      343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy      403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy      463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy      523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy      583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy      643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy      703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
            ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db      604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy      763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy      823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy      883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy      943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
            ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db      844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy     1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
            |||||| || |||||||||||||||||||||||||||||    ||||||||||||||||
Db      904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy     1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy     1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
            ||||                                          ||||||||||||||
Db     1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy     1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
            |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy     1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy    1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy    1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy    1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy    1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy    1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy    1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy    1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
           ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy    1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db    1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy    1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
           ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db    1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy    1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
           |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db    1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy    1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
           ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db    1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy    1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
           | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db    1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy    1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy    2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy    2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy    2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy    2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy    2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy    2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db    2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy    2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy    2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy    2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy    2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy    2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy    2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy    2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy    2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy    2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
           ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy    2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy    3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy    3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
           |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db    3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy    3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy    3181 GTCGGCGTAGGCTACCGGTTC 3201
           |||||||||||||||||||||
Db    3124 GTCGGCGTAGGCTACCGGTTC 3144


RESULT 21
AX236407
LOCUS       AX236407                4425 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 100 from Patent WO0164922.
ACCESSION   AX236407
VERSION     AX236407.1  GI:15796026
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4425)
```

```
  AUTHORS   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guiliani,M.M. and Pizza,M.C.
  TITLE     Heterologous expression of neisserial proteins
  JOURNAL   Patent: WO 0164922-A 100 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4425
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-ORF46.1"
BASE COUNT     1165 a   1259 c   1219 g    782 t
ORIGIN

  Query Match            92.8%;  Score 2971.8;  DB 6;  Length 4425;
  Best Local Similarity  97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy        103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy        163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy        223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy        283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy        343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy        403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy        463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy        523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy        583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy        643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy        703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
              ||||||||||||||||||||| |||||| ||||||||||||||| |||  ||||||||||
Db        604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy        763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy        823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy        883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy        943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy       1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
              |||||| || |||||||||||||||||||||||||||||    |||||||||||||||||
Db        904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy       1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy       1123 TTCA------------------------------------------ATGGCTCCAACCAT 1140
              ||||                                          ||||||||||||||
Db       1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy       1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy       1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy       1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db       1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy       1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
              ||||||  ||||||||||| |||| |||||||||||||||||||||||||||||||||||
Db       1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy       1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
              |||||||||||||||||||| | ||||||||||||||||||||||||||||||||| |||
Db       1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy       1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
              ||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db       1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy       1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
              | |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db       1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy       1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy     2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy     2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy     2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy     2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy     2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy     2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||  |||
Db     2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy     2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy     2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy     2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy     2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy     2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy     2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703

Qy     2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy     2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy     2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy     2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy     3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy     3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
              |||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||||
Db     3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy     3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy     3181 GTCGGCGTAGGCTACCGGTTC 3201
              |||||||||||||||||||||
Db     3124 GTCGGCGTAGGCTACCGGTTC 3144


RESULT 22
AX239715
LOCUS       AX239715                4425 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 17 from Patent WO0164920.
ACCESSION   AX239715
VERSION     AX239715.1  GI:15797369
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4425)
  AUTHORS   Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
            and Pizza,M.
  TITLE     Hybrid expression of neisserial proteins
  JOURNAL   Patent: WO 0164920-A 17 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..4425
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="deltaG983-ORF46.1"
BASE COUNT     1165 a   1259 c   1219 g    782 t
ORIGIN

  Query Match             92.8%;  Score 2971.8;  DB 6;  Length 4425;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches   47;  Indels   42;  Gaps    1;

Qy      103 ACTTCTGCGCCCGACTTCAATGCAGGCGGCACCGGTATCGGCAGCAACAGCAGAGCAACA 162
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        4 ACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACA 63

Qy      163 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       64 ACAGCGAAATCAGCAGCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGA 123

Qy      223 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 AGCATGCTCTGTGCCGGTCGGGATGACGTTGCGGTTACAGACAGGGATGCCAAAATCAAT 183

Qy      283 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTACAAACCCAAATGACGCATACAAGAAT 342
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      184 GCCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAT 243

Qy      343 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 TTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGT 303

Qy      403 ATCGTCGATACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 462
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      304 ATCGTCGACACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAA 363

Qy      463 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 GAACACGGCTATAACGAAAATTACAAAAACTATACGGCGTATATGCGGAAGGAAGCGCCT 423

Qy      523 GAAGACGGAGGCGGTAAAGACATTAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 582
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      424 GAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACT 483

Qy      583 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATGTGGTCTCC 642
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
```

```
Db    484 GAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCC 543

Qy    643 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 702
Db    544 CATATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCG 603

Qy    703 ACGCTACACATAATGAATACGCATGATGGAACCAAGAACGAAATAATGTCTGCAGCCATC 762
Db    604 ACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATC 663

Qy    763 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 822
Db    664 CGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGA 723

Qy    823 ACAACATCGAGGGCAGGCACTGCCGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTAC 882
Db    724 ACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTAC 783

Qy    883 CGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 942
Db    784 CGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATG 843

Qy    943 CAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATT 1002
Db    844 CAACAGAGCGATTACGGCAACCTGTCCTACCACATCCGTAATAAAAACATGCTTTTCATC 903

Qy   1003 TTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACACTGACCCTATTGCCATTTTAT 1062
Db    904 TTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT 963

Qy   1063 GAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1122
Db    964 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAG 1023

Qy   1123 TTCA-----------------------------------------ATGGCTCCAACCAT 1140
Db   1024 TTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCAT 1083

Qy   1141 TGCGGAATTACTGCCATGTGGTGCCTATCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1200
Db   1084 TGCGGAATTACTGCCATGTGGTGCCTGTCGGCACCCTATGAAGCAAGCGTCCGTTTCACC 1143

Qy   1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
Db   1144 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1203

Qy   1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
Db   1204 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG 1263

Qy   1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
Db   1264 CTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1323

Qy   1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
Db   1324 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1383

Qy   1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
Db   1384 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1443

Qy   1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
Db   1444 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1503

Qy   1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
Db   1504 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1563

Qy   1621 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGC 1680
Db   1564 GTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGC 1623

Qy   1681 GACGGCATTGTCTATCTGGCAGATACCGACCGATCCGGCGCAAACGAAACCGTGCACATC 1740
Db   1624 GACGGCATTGTCTATCTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATC 1683

Qy   1741 AAAGGCGATCTGCAGCTGGGCGGCGAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1800
Db   1684 AAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG 1743

Qy   1801 AAAGTGGACGGTACGGCGATGACCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAAGGG 1860
Db   1744 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAGGGG 1803

Qy   1861 GCAGGCTATCTCAACCGTACCGGACAACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1920
Db   1804 GCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGG 1863

Qy   1921 CGGGATTATTCTTTCTTCACAAACATCGAAACCGACGGTGGTCTGCTGGCTTCCCTCGAC 1980
Db   1864 CAGGATTATTCTTTCTTCACAAACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGAC 1923

Qy   1981 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 2040
Db   1924 AGCGTCGAAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCCTATTATGTCCGTCGCGGC 1983

Qy   2041 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2100
Db   1984 AATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCCGGTCTGAAACACGCC 2043

Qy   2101 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2160
Db   2044 GTAGAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCA 2103

Qy   2161 TCCGCAACACCCGAGACGGTTGAAACTGCGGCCGCCGACCGCACAGATATGCCGGGCATC 2220
Db   2104 TCCGCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATC 2163

Qy   2221 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2280
Db   2164 CGCCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAC 2223

Qy   2281 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2340
Db   2224 GGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCCAT 2283

Qy   2341 GCCGATATGCAGGGACGCCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGCTACG 2400
Db   2284 GCCGATATGCAGGGACGCCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCACG 2343

Qy   2401 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2460
Db   2344 GGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGAACAGGGCGGTGTT 2403

Qy   2461 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2520
Db   2404 GAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAAAACCGGCGAAAAT 2463

Qy   2521 ACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGAAAACAGTGCAAAT 2580
Db   2464 ACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGCGAAAACAGTGCAAAT 2523

Qy   2581 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2640
Db   2524 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGC 2583

Qy   2641 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2700
Db   2584 TATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGT 2643

Qy   2701 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2760
Db   2644 GCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGC 2703
```

```
Qy     2761 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2704 GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGGCGGTCTGCGCTAC 2763

Qy     2821 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2764 GACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAAC 2823

Qy     2881 AGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2940
            ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     2824 AGCCTCACTGAAGGCACGCTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGC 2883

Qy     2941 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2884 GATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTAC 2943

Qy     3001 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2944 ACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAAT 3003

Qy     3061 ATGCCGCACACCCGCCTGGTTGCCGGTCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3120
            |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db     3004 ATGCCGCACACCCGTCTGGTTGCCGGCCTGGGCGCGGATGTCGAATTCGGCAACGGCTGG 3063

Qy     3121 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3064 AACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGACGA 3123

Qy     3181 GTCGGCGTAGGCTACCGGTTC 3201
            |||||||||||||||||||||
Db     3124 GTCGGCGTAGGCTACCGGTTC 3144


RESULT 23
AF169448/c
LOCUS       AF169448                 600 bp    DNA     linear   BCT 09-AUG-2000
DEFINITION  Neisseria meningitidis strain Z2491 clone Cm024 unknown sequence.
ACCESSION   AF169448
VERSION     AF169448.1  GI:9754656
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Identification of regions of the chromosome of Neisseria
            meningitidis and Neisseria gonorrhoeae which are specific to
            pathogenic Neisseriae
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 600)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
            rue de Vaugirard, Paris 75015, France
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Neisseria meningitidis"
                     /strain="Z2491"
                     /db_xref="taxon:487"
                     /clone="Cm024"
BASE COUNT      111 a    159 c    144 g    184 t      2 others
ORIGIN

  Query Match             18.0%;  Score 576.4;  DB 1;  Length 600;
  Best Local Similarity   98.5%;  Pred. No. 1.1e-121;
  Matches  591;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy      306 CGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAAT 365
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      600 CGGAGACTTTACAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCTGCAAT 541

Qy      366 TGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGATACAGGCGAATCCGT 425
            ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||  | 
Db      540 TGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATNAGN 481

Qy      426 CGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTA 485
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      480 CGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTA 421

Qy      486 CAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACAT 545
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db      420 CAAAACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGACGGTAAAGACAT 361

Qy      546 TAAAGCTTCTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCG 605
            | |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db      360 TGAAGCTTCTTTCGACGACGAGGCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCG 301

Qy      606 CCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGT 665
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      300 CCACGTAAAAGAAATCGGACACATCGATGTGGTCTCCCATATTATTGGCGGGCGTTCCGT 241

Qy      666 GGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCA 725
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      240 GGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGCA 181

Qy      726 TGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGG 785
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      180 TGATGGAACCAAGAACGAAATAATGTCTGCAGCCATCCGCAATGCATGGGTCAAGCTGGG 121

Qy      786 CGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGC 845
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      120 CGAACGTGGCGTGCGCATCGTCAATAACAGTTTTGGAACAACATCGAGGGCAGGCACTGC 61

Qy      846 CGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTC 905
            ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||| |
Db       60 CGACCATTTCCAAATAGCCAATTCGGAGGAGCAGTACCG-CAAGCGTTGCTCGCCTATCC 2


RESULT 24
AF169473
LOCUS       AF169473                 286 bp    DNA     linear   BCT 09-AUG-2000
DEFINITION  Neisseria meningitidis strain Z2491 clone Em085 unknown sequence.
ACCESSION   AF169473
VERSION     AF169473.1  GI:9754681
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 286)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Identification of regions of the chromosome of Neisseria
            meningitidis and Neisseria gonorrhoeae which are specific to
            pathogenic Neisseriae
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 286)
  AUTHORS   Perrin,A., Nassif,X. and Tinsley,C.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156
            rue de Vaugirard, Paris 75015, France
FEATURES             Location/Qualifiers
     source          1. .286
                     /organism="Neisseria meningitidis"
                     /strain="Z2491"
                     /db_xref="taxon:487"
                     /clone="Em085"
BASE COUNT       88 a     73 c     63 g     62 t
ORIGIN

  Query Match              8.5%;  Score 270.8;  DB 1;  Length 286;
  Best Local Similarity   99.0%;  Pred. No. 2e-51;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy      925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy      985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACAC-ACT 1043
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACAACT 180

Qy     1044 GACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGT 1103
            ||||||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||
Db      181 GACCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTTGCAGGCGT 240

Qy     1104 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 1149
            ||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 286


RESULT 25
A68924
LOCUS       A68924                   286 bp    DNA     linear   PAT 06-MAY-1999
DEFINITION  Sequence 95 from Patent WO9802547.
ACCESSION   A68924
VERSION     A68924.1  GI:4759843
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 286)
  AUTHORS   Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and
            Merker,P.
  TITLE     DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS
            SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL
            APPLICATIONS
  JOURNAL   Patent: WO 9802547-A 95 22-JAN-1998;
            INST NAT SANTE RECH MED (FR)
COMMENT     Other publication FR 2751000 19980116.
FEATURES             Location/Qualifiers
     source          1..286
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       88 a     73 c     63 g     62 t
ORIGIN

  Query Match              8.5%;  Score 270.8;  DB 6;  Length 286;
  Best Local Similarity   99.0%;  Pred. No. 2e-51;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      865 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 AATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60

Qy      925 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 984
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACTTGTCCTACCACATCCGTAAT 120

Qy      985 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACAC-ACT 1043
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      121 AAAAACATGCTTTTCATTTTTTCGGCAAGCAATGACGCACAAGCTCAGCCCAACACAACT 180

Qy     1044 GACCCTATTGCCATTTTATGAAAAAGATGCTCAAAAAGGCATTATCACAGTCGCAGGCGT 1103
            ||||||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||
Db      181 GACCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTTGCAGGCGT 240

Qy     1104 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 1149
            ||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGAATT 286


RESULT 26
BPE318229
LOCUS       BPE318229               3493 bp    DNA     linear   BCT 10-JUL-2001
DEFINITION  Bordetella pertussis sphB1 gene for autotransporter subtilisin-like
            protease.
ACCESSION   AJ318229
VERSION     AJ318229.1  GI:14715184
KEYWORDS    autotransporter subtilisin-like protease; SphB1 protein.
SOURCE      Bordetella pertussis.
  ORGANISM  Bordetella pertussis
            Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
            Bordetella.
REFERENCE   1  (bases 1 to 3493)
  AUTHORS   Jacob-Dubuisson,F. and Coutte,L.
  TITLE     An autotransporter subtilisin-like protein serves as a maturation
            protease in the secretion pathway of a Gram- negative bacterium
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3493)
  AUTHORS   Jacob-Dubuisson,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-2001) Jacob-Dubuisson F., Institut de Biologie de
            Lille, Institut Pasteur de Lille, 1 rue Calmette, 59019 Lille
            Cedex, FRANCE
FEATURES             Location/Qualifiers
     source          1..3493
                     /organism="Bordetella pertussis"
                     /isolate="Tohama I"
                     /db_xref="taxon:520"
     gene            165..3284
                     /gene="sphB1"
     CDS             165..3284
                     /gene="sphB1"
                     /function="proteolytic maturation of FhaB"
                     /codon_start=1
                     /transl_table=11
                     /product="autotransporter subtilisin-like protease"
                     /protein_id="CAC44081.1"
                     /db_xref="GI:14715185"
                     /translation="MPAQRTPRTAVCEATVRSSPRWIHCTGFVLCALLAACGGGGGG
                     GGGGGGGSPGGRAPSAPQPAPSPRPEPAPEPAPNPAPRPAPQPPAPAPGAPRPPAPPP
                     EAPPPVMPPPAVPPQLPEVPAADLPRVRAPLSTYRRPQRTDFVTPTGGPFFAKQDKAL
                     NTIDLKMAHDLKLRGYRVKVAVVDEGVRSDHPLLNVEKKYGGDYMADGTRTYPDPKRQ
                     GRHGTSVALVLAGQDTDTYRGGVAPNADLYSANIGTRAGHVSDEAAFHAWNDLLGHGI
                     KIFNNSFATEGPEGEQRVKEDRNEYHSAANKQNTYIGRLDRLVRDGALLIFAAGNGRP
                     SGRAYSEVGSVGRTPRVEPHLQRGLIVVTAVDENGRLETWANRCGQAQQWCLAAPSTA
                     YLPGLDKDNPDSIHVEQGTSLSAPLVTGAAVLVQDRFRWMDNDNLRTTLLTTAQDKGP
                     YGVDPQYGWGVLDVGRAVQGPAQFAFGDFVARVTDTSTFGNDISGAGGLVVDGPGALV
                     LAGSNTYAGRTTIKRGTLDVFGSVTSAVTVEPGGTLTGIGTVGTVTNQGTVVNKEAGL
                     HVKGDYSQTAQGLLVTDIGSLLDVSGRASLAGRLHVDDIRPGYVGGDGKSVPVIKAGA
                     VSGVFATLTRSPGLLNARLDYRPQAVYLTMRRAERVHAAAQRGADDGRRASVLAVAE
                     RLDAAMRELDALPESQRDAAAPAAAIGRIQRVQSRKVLQDNLYSLAGATYANAAAVNT
                     LEQNRWMDRLENHLAQAGGERVAAIAEYRHGQLRWRPDGLQGRQRGNGIMLGLAREVS
                     AGLSLAAALTHSRTHWDESSGAPARDNAAMTTPGVLLGARRAWEDGWFVQGALGYSRY
                     RNQATRHISLGDAGHTVGATARGHVWQADAGLGRQWTLAPGHTLAPRAGLQLTHLRQQ
                     GFSESGAQGLGLRAHALTRTVPTLWAQLQSRHAFMLGATPMTAQLQLGVWHDLRARRY
                     AASGGFAGLAQDQGASGYWPVPRTRVQGALGLRAEFAPGLVLGLGYTGQLATHWVDHQ
                     LSASLTYRY"
BASE COUNT      555 a   1343 c   1138 g    457 t
ORIGIN

  Query Match              4.6%;  Score 147.8;  DB 1;  Length 3493;
  Best Local Similarity   61.0%;  Pred. No. 4.5e-23;
  Matches  264;  Conservative    0;  Mismatches  157;  Indels   12;  Gaps    1;

Qy     1225 GGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATAC 1284
            || || ||  | ||||| ||  | || |||||| | || |  ||| ||| | |    |
Db     1392 GGCACGTCGCTATCCGCGCCGCTGGTCACCGGCGCCGCCGTACTGGTGCAGGATCGCTTT 1451

Qy     1285 CCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGT 1344
            |  ||||||  ||||||||||||||| |||||  ||||||| ||||| |||||||  ||
Db     1452 CGCTGGATGGACAACGACAACCTGCGCACCACCTTGCTGACCACGGCGCAGGACAAGGGC 1511

Qy     1345 GCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAAC 1404
             |   |||||| |||    ||| |||||||||  |||| || | |||    ||| || |
Db     1512 CCGTACGGCGTCGACCCGCAGTACGGCTGGGGCGTGCTCGACGTGGGCCGCGCCGTGCAG 1571
```

```
Qy    1405 GGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATT 1464
              || || |||    ||  | ||||||||||   ||||        |  | ||
Db    1572 GGCCCGGCGCAGTTCGCCTTCGGCGACTTCGTCGCCCGGGTTACGGATA----------- 1620

Qy    1465 GCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGC 1524
               ||| | | ||| | |||||||| || ||| | ||||| ||| ||      | |||  | ||
Db    1621 -CCTCCACGTTCGGCAACGACATCTCCGGCGCCGGCGGGCTGGTCGTCGACGGCCCCGGC 1679

Qy    1525 CAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCG 1584
                |||   |||  ||||  ||| |||||| | |||     ||||   ||| |   ||||   |
Db    1680 GCGCTGGTCCTGGCCGGCTCCAATACCTATGCCGGCCGCACCACCATCAAGCGCGGCACC 1739

Qy    1585 CTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCTGATT 1644
              || |    | |  |||| |   ||   || |    |    ||||||  |       ||  ||||||
Db    1740 CTCGACGTCTTTGGCAGCGTCACGTCCGCCGTCACCGTCGAGCCTGGCGGCACGCTGACC 1799

Qy    1645 TATAACGGGGCGG 1657
                 | |||   | |
Db    1800 GGCATCGGCACCG 1812
```

```
RESULT 27
AE011850
LOCUS       AE011850              11292 bp    DNA     linear   BCT 29-MAY-2002
DEFINITION  Xanthomonas axonopodis pv. citri str. 306,  section 228 of 469 of
            the complete genome.
ACCESSION   AE011850 AE008923
VERSION     AE011850.1  GI:21108371
KEYWORDS    .
SOURCE      Xanthomonas axonopodis pv. citri str. 306.
  ORGANISM  Xanthomonas axonopodis pv. citri str. 306
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xanthomonas.
REFERENCE   1  (bases 1 to 11292)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Comparison of the genomes of two Xanthomonas pathogens with
            differing host specificities
  JOURNAL   Nature 417 (6887), 459-463 (2002)
  MEDLINE   22022145
   PUBMED   12024217
REFERENCE   2  (bases 1 to 11292)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-NOV-2001) Departmento de Bioquimica, Universidade de
            Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
            Brazil
FEATURES             Location/Qualifiers
     source          1..11292
                     /organism="Xanthomonas axonopodis pv. citri str. 306"
                     /strain="306"
                     /db_xref="taxon:190486"
                     /note="pathovar: citri"
     gene            456..3191
                     /gene="XAC2144"
     CDS             456..3191
                     /gene="XAC2144"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="serine protease"
                     /protein_id="AAM36997.1"
                     /db_xref="GI:21108372"
                     /translation="MNRSLLATTIASCLILTACGGGGGGNVRSDAPQTTVIAPTPTT
                     PPPAADPLPAPAPAPTPAPPRYQGVGSNLLVPTNADLAQQAGARGQGVKLAVLDDNLV
                     PSYAPISGKVDSFNDYTASPGTPESSANALRGHGTIVSALVLGSAQDGFAGGVAPDAD
                     LFYARICAENSCGTQQTRRAAVDLAAAGVRIANLSIGASYPDAAASANAALAWKYALT
                     PLVQADALIVASTGNEGAAEASYPAATPVQEASVRNNWLAVGAINIDSAGNAAGLTSY
                     SNHCGAAAQWCLVAPGSYTAPALAGTELQGQIAGTSFSTAAVSGVAAQVLGVYPWMSA
                     SNLQQTLLTTATDLGDPGVDALYGWGLVNAAKAIKGPGQFASNWAANVTAGYDSTFSN
                     DISGAGSLTKSGAGKLTLSGTNTYSGGTTIQDGVLSLTGSLRSDLLVANSATFESRGG
                     VVNGNYSLVNATGTTAIELGKGLAVTGQASLKGNLLLLPEAAGYTVGSTEKIVKAGVL
                     QGTFDRVQYANNFFWTAALSYDPTTVTATLTRNASASASAQAAGAPQSVVNGANQADAL
                     VKVLDTRVASGDTDNLGGLISATGSLIAASDAQVAASLPTLTGQVHGIERTLGFQSAL
                     NDARVAADRLPYLADTTTLTTWMQGSYLDGDLGRSGFDSADYTQQATTFGVDLPMGNA
                     GTIVGAAVTSGQDRAHIAASSSRLRSDRYAMTAYLYQPIGRAYLSATGSYGMSNVETE
                     RNVQAGSIGERLQDRRHDTNWSARAEIGLMPQAGLSPFVAAGVVSQTQDAFSEASVTG
                     LGLSANSDTLRAGYGEVGVRGHLARGKWGFDGLVAYRSLLGNPNSDFSAWFTGLPEAR
                     FNVAGEPVSKQSLRASVGLRYQLNDAFSIYGNVGAERGRSDTNSINANVGLRWQF"
     gene            complement(3157..3546)
                     /gene="XAC2146"
     CDS             complement(3157..3546)
                     /gene="XAC2146"
                     /note="putative; ORF located using Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAM36998.1"
                     /db_xref="GI:21108373"
                     /translation="MAHRHRCEVARQPSGLGQLQRVLQPRRWRAWCQPSVRETRGFWA
                     FSPVHGHRIDIPSYRLRAVGTQDAVSAKRLAVRHSHAAVNEKAPDMPGPFFYLSRSRG
                     RDIITCVTCSTAASAAIRTASAGRRWH"
     gene            3545..4795
                     /gene="czcB"
                     /note="XAC2145"
     CDS             3545..4795
                     /gene="czcB"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="cation efflux system protein"
                     /protein_id="AAM36999.1"
                     /db_xref="GI:21108374"
                     /translation="MGARRCRWTPLPAASVDPSPMQPRPAFAHRRPRVQGLPSPRSSH
                     RMPSPFPRSTLAAACVCSLLAGCGATPPPAAPSPLLIHTNMQVMVPAHSPLRASLRVQ
                     AVRSQPIAPLLQMPAVIEVLPERLVRVVPPLAGRVVALPKTLGDTVRAGDVLCVLDSA
                     ELASAYSDAGKARATLEQARLELARQKALAADSISAARDLQAAQQAFDSAQNDARAAS
                     DRLAQLGVAAQASSHRRYVVRAPIAGRLVDLSAALGGFWNDTSASLMTVADISQVWLT
                     ASVPEREIGQVFEGQPVTASLEAYPAQRFVGQVQHLDDLLDPATRTLKVRVALENRDG
                     LLKPGMFARAQFHSRPRQALLVPDSALLQMGLYTRVFIETAPFAYRSRIVATGRAMDG
                     YTEILSGLKAGERVVIKDGALLND"
     gene            4788..7880
                     /gene="czcA"
                     /note="XAC2147"
     CDS             4788..7880
```

```
                     /gene="czcA"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="cation efflux system protein"
                     /protein_id="AAM37000.1"
                     /db_xref="GI:21108375"
                     /translation="MIEAITTTCFQRRGIVWLVFVLIALYGTWSWTQLPVEAYPDIAD
                     VTSQVVTQVPGLGAEEVEQQITVPLERALMGTPGLHVLRSRSLFALSLITLVFDDGTE
                     GYFARQRVLERIQAVTLPYGAIPGLDPYTSPTGEIYRYTLESKTRSLRELSDLQFWTV
                     IPRLQKVPGVADVTNFGGLTTQFSLALEPDRLTRYGVSLQQVKSAITSNNADGGGSVM
                     DRGEQSYVIRGIGLLHSLQDIGNVVVSSGNGVPVLVKDLGEVRYDNVERRGILGKDGN
                     PDTIEGIALLLKDSNPSVALQGIHSAVEELNNSALPKDVKVVPYLDRTALIDATMHTV
                     SATLTEGMLLVCVVLLIFLGSPRAAAIVSLTIPLSLLIAFIVMHHLKIPANLLSLGAI
                     DFGILVDGAVVLVENVLRLREENSQRALTARDAIDATLQVARPIFFGMAVIGCAYLPL
                     LAFERIEYKLFSPMAYAVGAALIGALLVALALIPGLAWLAFRKPRKMLHNRVLEELGQ
                     RYRAVLERSVGRRGWLLACAALALCVLALLGGSIGRDFLPYIDEGSLWLQVQMPPGIT
                     LDKAATMANALRKATLEFPEVSYVVTQTGRNDDGTDYWTPSHIEASVGLRPYKEWPAG
                     MDKQALIAALGARYARMPGYTVSMMQPMIDGVQDKLSGAHSDLTVKVFGDDLQQVRGV
                     ADQVAFALHKVPGAADIAVDVEPPLPNLQVRFDREAAARYGINAADVSDLISTGIGGS
                     PIGQMYLGQKSYDLTVRFPQRYRNDPQAIGALRLRTAAGAEIPLSAVANIATTSGQSV
                     IVREMGRRNIIVRLNVRGRDLSSFLSDAQATLARQVHLDPQHMQLVWGGQFENLQRAQ
                     ARLLVVLPTTLCIMFVLLFGAFGNLRQPTLVLAAVPLAMIGGLAALHLRGMTLNVSSA
                     VGFIALFGVAVLNAVLMLAQINRLRQDVGMSLREAVVAGAVSRMRPVLMTATVAALGL
                     APAMLATGLGSDVQRPLATVVVGGLVTATLLTLVLLPSLYYLMEAFVATRQGRQGTTP
                     "
     gene            7877..9361
                     /gene="ttgC"
                     /note="XAC2148"
     CDS             7877..9361
                     /gene="ttgC"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="outer membrane efflux protein"
                     /protein_id="AAM37001.1"
                     /db_xref="GI:21108376"
                     /translation="MTDVRNALLASAVAASLAACAVGPDFQRPAAPAADRYTANALPV
                     TTAAADVAHGGAQEFEPGQELPAAWWRAFGSPQLNAVVERALRANPDLQAADAALRQA
                     QQTLAAQRGAWRPQADLRVDASRQRDSVVPAPDAGIDTPYALRTAQLSVSYTLDVFGG
                     TRRQVEAAGAQAQAQRDQRDAAYLSLTANVVNAAIGEAAVRAQLDAARAQVAIAKQLC
                     TLTARQQALGALGAAQVLAQRTALAQAQAAVPALEKQLEQQRTLLAILGGQLPGDAAG
                     EAFTLDTLQLPAQLPLSLPSRLLDRRPDIRAAEAQAHAASALIGVAAAARLPSISLSA
                     ASGGSASGVGRMFAAGNTQWSVAGGLLQPLFHGGALAHQQRAAEAGYTQAMAQYRSVV
                     LNAFGEVSNALTALQSDAQSLHACADAQALAERTLALVRRQQAVGAVGMADLLQAQQA
                     YQSTQAATVQAQAARYTDTVALYQALGGGPWWSPAARTPKPAVP"
     gene            9129..10502
                     /gene="nodV"
                     /note="XAC2149"
     CDS             9129..10502
                     /gene="nodV"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="nodulation protein"
                     /protein_id="AAM37002.1"
                     /db_xref="GI:21108377"
                     /translation="MRKRWRSARWRWCDVNRRLVRWAWRICCRRSRPTRARRPPRSRH
                     RPRATPTPWRCIRPWVAGRGGRQRPARPSPRFHESRAASSAGCVTAAPPAETIAAMRM
                     TSLFRSPPTRTDLLGMVGLVAGMAAIFVADTFTDYAVAAALFYTLVVLLSVRLLTRTG
                     VVLLAGTCLVLIVLSFHLTPAGSYRVGVINSCISAVVVAVTTYLAVQMEAAKASAHAA
                     QARLLRLTRSGSMGGLASSIAHEVNQPLAAIVTSGNACQRWLAQQPPNLAKAAAALER
                     MLADAGRASAIIARMRALARGEASQRQPLDLNRTVLDMIALTRGETERHGIALSLELD
                     PALPAVLADGVQIQQVIGNLLLNAVEALATLPDEARSIQIRSHIDGAWACLCIDDTGP
                     GLSDDALAHVFEAFWSTKDDGMGIGLSISRTIMEANGGQIWAEPAGVRGARFGLRFPL
                     VATERRQ"
     gene            10499..11110
                     /gene="nodW"
                     /note="XAC2150"
     CDS             10499..11110
                     /gene="nodW"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1

  Query Match             4.1%;  Score 132.4;  DB 1;  Length 11292;
  Best Local Similarity   55.7%;  Pred. No. 1.7e-19;
  Matches  303;  Conservative    0;  Mismatches  226;  Indels   15;  Gaps    2;

Qy       1087 ATCACAGTCGCAGGCGTAGACCGCAGTGGAGAAAAGTTCAATGGCTCCAACCATTGCGGA 1146
Db       1239 ATCAATATCGATAGCGCCGGCAATGCCGCCGGCCTGACCAGCTATTCCAATCATTGCGGC 1298

Qy       1147 ATTACTGCCATGTGGTGCCTATCGGCACC------CTATGAAGCAAGCGTCCGTTTCACC 1200
Db       1299 GCCGCTGCGCAGTGGTGCCTGGTCGCGCCGGGCAGCTATACCGCCCCGGCCCTGGCCGGC 1358

Qy       1201 CGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACG 1260
Db       1359 ACCGAATTGCAGGGCCAGATTGCCGGCACCTCGTTCTCCACCGCCGCCGTCAGCGGTGTC 1418

Qy       1261 GCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTG 1320
Db       1419 GCTGCGCAGGTGCTGGGCGTGTATCCGTGGATGAGCGCTTCCAACCTGCAGCAAACCCTG 1478

Qy       1321 CTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTG 1380
Db       1479 CTGACCACCGCCACCGACCTGGGCGACCCGGGCGTGGACGCGCTGTACGGCTGGGGCCTG 1538

Qy       1381 CTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCC 1440
Db       1539 GTCAATGCGGCCAAGGCGATCAAGGGCCCGGGCCAGTT---------CGCAAGCAACTGG 1589

Qy       1441 GATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGC 1500
Db       1590 GCGGCCAATGTGACCGCTGGCTACGACAGCACCTTCAGCAACGACATCTCCGGCGCGGGC 1649

Qy       1501 GGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAACACCTATACGGGC 1560
Db       1650 AGCCTGACCAAGAGCGGCGCAGGCAAGCTCACCCTGAGCGGCACCAATACCTATAGCGGC 1709

Qy       1561 AAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGC 1620
Db       1710 GGCACCACGATTCAGGATGGCGTGCTGTCGCTCACCGGCAGCCTGCGTTCGGATCTTCTG 1769

Qy       1621 GTCG 1624
Db       1770 GTGG 1773


RESULT 28
AE012308/c
LOCUS       AE012308               11019 bp    DNA     linear   BCT 23-MAY-2002
DEFINITION  Xanthomonas campestris pv. campestris str. ATCC 33913,  section 216
            of 460 of the complete genome.
ACCESSION   AE012308 AE008922
VERSION     AE012308.1  GI:21113147
KEYWORDS    .
SOURCE      Xanthomonas campestris pv. campestris str. ATCC 33913.
  ORGANISM  Xanthomonas campestris pv. campestris str. ATCC 33913
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xanthomonas.
REFERENCE   1  (bases 1 to 11019)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
```

```
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Comparison of the genomes of two Xanthomonas pathogens with
            differing host specificities
  JOURNAL   Nature 417 (6887), 459-463 (2002)
  MEDLINE   22022145
   PUBMED   12024217
REFERENCE   2  (bases 1 to 11019)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-NOV-2001) Departmento de Bioquimica, Universidade de
            Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
            Brazil
FEATURES             Location/Qualifiers
     source          1. .11019
                     /organism="Xanthomonas campestris pv. campestris str. ATCC
                     33913"
                     /strain="ATCC 33913"
                     /db_xref="ATCC:33913"
                     /db_xref="taxon:190485"
                     /note="pathovar: campestris"
     gene            complement(73. .2898)
                     /gene="XCC2025"
     CDS             complement(73. .2898)
                     /gene="XCC2025"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="serine protease"
                     /protein_id="AAM41314.1"
                     /db_xref="GI:21113148"
                     /translation="MKRSLLATTISSCLVLTACGGGGGGGNVRNDPPPTTTVVPAPIT
                     DPTPAPAPTPAPTPAPTPAPTPSPAPAPAPAPPPTADPLPKTAPVPVAPPRYQGLASN
                     LLVPTNADLAQQAGARGQGVKIGVMDDNLVQTYAPVAGKVDAFTDYTAVPGAAESTSN
                     RLRGHGSVVSALVLGSAQDGFAGGVAPDADLIYGRICAENSCGTQQARRAAVDMAAAG
                     VRIANLSIGASYADAASSANAALAWRFALTPLVQADALIVASTGNDGAAEASYPAAAP
                     VQEASLRNNWLAVGAVEIDSAGNPAGLSSYSNHCGSAAQWCLVAPGMYAAPALAGTEL
                     QGQIAGTSFSTAAVSGVAAQVLGVYPWMSASNLQQTLLTTATDLGDPGVDALYGWGMV
                     NAAKAIKGPGQFASNWAANVTSGYDSTFSNDISGSGNLTKNGAGRLTLSGNNSYAGGT
                     SISDGVLALSGSVRSDVLVLNSATFESRGGTINGNYSLVNSTGTTAIELGKGLTVTGQ
                     ASLKGNLLLLPEAAGYTVAGSERIVRAGALQGTFDRVQYANNFFWTAALAYDPTSVTA
                     TLTRNAAAASAQAAGASQSVVAGARQADSLVKALDTRVASGDTANLGGVITATGALLL
                     ATDAQVASSLPTLVGQVHGIERTLGFQSALNDARVAADRLPLLASTTTLTTWMQGSYL
                     DGDLSRAGYASADYTQRATTVGFDMPLGGAGTIAGAALTSGEDRGTISASASRLRADR
                     YAMTAYLYQPVGMAYLSAIGSYGMGSVETERSVQAGSIGERLQDRRHDTNWSLRTELG
                     LLPRAGLSPFVAAGVVSQTQDAFSEASVTGLGLSAASDTLRAGYGEVGVRGHLAHGKW
                     GFDGLVAYRSLFGDNRADFTAWFTGLPEARFTVAGEPLSKEALRASVGVRYSVSDTFS
                     LYGNVGAERGSSDASSVNANVGLRWQF"
     gene            complement(4402. .5037)
                     /gene="XCC2027"
     CDS             complement(4402. .5037)
                     /gene="XCC2027"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="transcriptional regulator"
                     /protein_id="AAM41315.1"
                     /db_xref="GI:21113149"
                     /translation="MPRLPISAIMSCAMAAIVSGAVPPEPPMPRAHIVETAAALFRER
                     GYDGVGVAELMAAAGFTHGGFYKHFRSKADLMAETAALGFSKTAAASDAVDVAEFLSD
                     YVSRKHRDSRARGCTMAALCGDAARQPEAIKAEFASGISTLLDTLANADDARNAAQRQ
                     QVRADTIDRLAHAIGAVVLSRACPDDSPLADEILEVCRRKILAALPGQPAR"
     gene            4997. .5590
                     /gene="W78"
                     /note="XCC2026"
     CDS             4997. .5590
                     /gene="W78"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="nodulation related protein"
                     /protein_id="AAM41316.1"
                     /db_xref="GI:21113150"
                     /translation="MAHDMIALIGSLGIAQVDLLGTSFGGFVAQDIVRQAPQCVRKLI
                     LASSGPAGGAGIDKVGAISLPQILKGALTRRDPKYFLFSTASQRGRHAATAFLDRLAE
                     RTQHRDKPLSPRAFWRQLQAIKAWARQPAQDLTNFTMPVLITAGDHDRLVPPVHSEAL
                     ARRIRHAQLIVYQDAGHGAAFQYHAAFVASVLAFLDR"
     gene            complement(5807. .7771)
                     /gene="XCC2028"
     CDS             complement(5807. .7771)
                     /gene="XCC2028"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM41317.1"
                     /db_xref="GI:21113151"
                     /translation="MRHRRSFRIEAVLTAAALAAAVLPGAAAAESATPATAPAVFPAP
                     TSLRLLGPELSLGTSVVLVRPKGTDAASEALVRKVLTTAGVRKIRVATTPASAPGIVQ
                     VVLGLSETASVRSALEQTGAALPERDEAYALGSRASGSGVTIVLAGKDSDGLYHAAQT
                     FRQLVTAGQMPSVAIADAPAMPVRGSIEGFYGKPWSMAEREDHLAFLARFKANTYIYS
                     PKDDVFARERWREPYPAATLKALGKVVEVANREHINFVYALSPGPSVCYSDPAELDAI
                     RRKFSALRKRGVRSFYVAFDDIEYTKWNCKADEAAFGPSGEQAAATAQAKLLNAVQAD
                     IAAAGHGSLIMVPTEYFNATVSPYKTTLHKALDPRVVIQWTGTDVVPASISVTDARNA
                     TKAFGRKTLLWDNYPVNDFAETTGRLLLAPYDRRQAGLSAELSGIVSNPMNQEVASRP
                     AVAGLLAFAWNDVGYDAQRTWRYAARDLAGNDPVVTQALLTFFDTQHVAPTFGHLPWQ
                     PQAPRLKALVDDARDALASGDQAAITPALHALASAADALAAAPAQIRAGSTTQAFVRE
                     AAPWLTAAQLWGQSLRRSVDGVSAALAANSQAATQQFDAAKQLAEQAAAIQSIPGATR
                     FAGPVKVADGVLDSFVKDAPGLVYVPTAQE"
     gene            complement(8484. .9257)
                     /gene="XCC2029"
     CDS             complement(8484. .9257)
                     /gene="XCC2029"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM41318.1"
                     /db_xref="GI:21113152"
                     /translation="MPLRSVRTTLAAFVLVALAPVATAATQDEEAAVRAADAAFWKAY
                     NACDLPAAGKLLSADVEFYHDRTGLTRTRDGLIESLRKGPCADPNMRLRREAIDTSLA
                     FHPLSGGFALLSGQHRFYVQRAGAPESLDGQASFTTVWIADGGQWRMHRVLSYSHGPA
                     PYMPPARTLTLPAATLARYAGQYRSARIGSIVVVADGDHLRLTAGSFVATLYPESTTR
                     FFAMERDLRFDFEADAAGAIVALAAYEHGEVSDRAVREP"
     gene            complement(9387. .10658)
                     /gene="XCC2030"
     CDS             complement(9387. .10658)
                     /gene="XCC2030"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
```

```
                 /product="sensor histidine kinase"
                 /protein_id="AAM41319.1"
                 /db_xref="GI:21113153"
                 /translation="MRPSSAPSALNAIAEIERISAVPQILETVAHITGSRFTAVARVT
                 DTSWVACATYDTLGFGLKPGGQLEVESTICHEIRLSPKPVIFSHASQHPVYAHHHTPK
                 LYGLESYVSVPIFRRDGGFFGTLCAIDSAPANFDEAHILKSLEFFALLIGNSLETEEK
                 LQQAATALSAANDAGELRDQFIAVLGHDLRSPLQSISMATEMLQLEPQSPRGASLLGH
                 IRTSLTRIGGLVDDVLDFARGRLGGIPVTLRVEDALERRLLEVVQEVCSATGRTDLH
                 ADIHIGGEIVCDAPRLSQLFANLLSNAAVHGTPSIPVTLRAWQAAGILHLSVHNGGVI
                 APEKRAKLFSPFSRDDDDAPQPGLGLGLYIAAEIAKAHEGRLSVSSDLLAGTTFTFEM
                 PVAPAQPQHSVHAGDEDALARDDARPRAMQG"
BASE COUNT     1843 a   3451 c   3851 g   1874 t
ORIGIN

  Query Match             4.1%;  Score 130.2;  DB 1;  Length 11019;
  Best Local Similarity   57.1%;  Pred. No. 5.4e-19;
  Matches  284;  Conservative    0;  Mismatches  198;  Indels   15;  Gaps    2;

Qy       1132 TCCAACCATTGCGGAATTACTGCCATGTGGTGCCTATCGGCACC------CTATGAAGCA 1185
              ||||| ||||||||      ||||   |||||||||  ||||||        || |   ||
Db       1980 TCCAATCATTGCGGGTCCGCTGCGCAGTGGTGCCTGGTGGCACCGGGCATGTACGCCGCC 1921

Qy       1186 AGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCCGCACCC 1245
                 | ||     |  |              | |    || || ||||| || || || ||| |   ||
Db       1920 CCGGCCCTGGCCGGCACCGAGTTGCAGGGCCAGATCGCCGGCACCTCGTTCTCCACTGCC 1861

Qy       1246 ATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC 1305
                 || | ||||  ||| || | | | ||          || || ||||||||      ||||
Db       1860 GCTGTCAGCGGCGTGGCCGCGCAGGTACTCGGCGTGTATCCCTGGATGAGCGCTTCCAAC 1801

Qy       1306 CTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAG 1365
              ||||      || |||||||| || ||          ||| | || |       |||||| |||           |
Db       1800 CTGCAGCAGACCCTGCTGACCACCGCCACCGACCTGGGCGATCCCGGCGTCGACGCGTTG 1741

Qy       1366 TTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTC 1425
               | ||| |||||  || |  |||| |  ||||| || || || || || |        ||
Db       1740 TACGGTTGGGGCATGGTCAATGCCGCCAAGGCGATCAAGGGCCCGGGCCAGTT------- 1688

Qy       1426 GGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGAC 1485
                 || |    | |     |    | || |  || ||||        | |   | ||||| | ||||||
Db       1687 --CGCCAGCAACTGGGCGGCAAACGTGACCTCCGGCTACGACAGCACCTTCAGCAACGAC 1630

Qy       1486 ATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAAC 1545
              || || ||||  |||   ||||| ||| || |||| | ||  |||      ||| |||||||
Db       1629 ATCTCCGGCAGCGGCAACCTGACCAAGAACGGCGCCGGCAGACTCACCCTGAGCGGCAAC 1570

Qy       1546 AACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAAC 1605
              |||| |||  | |||     ||||  |||    | |||    |||||  ||||  |||| |
Db       1569 AACAGCTACGCCGGCGGCACCAGCATCAGCGACGGCGTGCTGGCCTTGTCGGGCAGCGTG 1510

Qy       1606 AAATCGGATATGCGCGT 1622
                 || ||  |||  ||
Db       1509 CGCTCCGACGTGCTGGT 1493


RESULT 29
NGOPAK
LOCUS       NGOPAK                  1434 bp    DNA     linear   BCT 10-APR-1995
DEFINITION  N.gonorrhoeae opaK gene for opacity protein.
ACCESSION   X52364 S36068 S75784
VERSION     X52364.1  GI:48702
KEYWORDS    opacity protein; opaK gene.
SOURCE      Neisseria gonorrhoeae.
  ORGANISM  Neisseria gonorrhoeae
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1434)
  AUTHORS   Bhat,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-1990) Bhat K., National Institutes of Health,
            Rocky Mountain Laboratory, Hamilton 59840, U S A
REFERENCE   2  (bases 1 to 1318)
  AUTHORS   Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
            Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
  TITLE     The opacity proteins of Neisseria gonorrhoeae strain MS11 are
            encoded by a family of 11 complete genes
  JOURNAL   Mol. Microbiol. 5 (8), 1889-1901 (1991)
  MEDLINE   92114767
   PUBMED   1815562
REFERENCE   3  (bases 1 to 1434)
  AUTHORS   Bhat,K.S., Gibbs,C.P., Barrera,O., Morrison,S.G., Jahnig,F.,
            Stern,A., Kupsch,E.M., Meyer,T.F. and Swanson,J.
  TITLE     The opacity proteins of Neisseria gonorrhoeae strain MS11 are
            encoded by a family of 11 complete genes
  JOURNAL   Mol. Microbiol. 6 (8), 1073-1076 (1992)
  MEDLINE   92261323
   PUBMED   1584024
COMMENT     The cloned gene is in its non-expressed state. The change in number
            of the coding repeat unit has caused the atg start codon in the
            signal peptide to be out of frame with the rest of its coding
            sequence.
FEATURES             Location/Qualifiers
     source          1..1434
                     /organism="Neisseria gonorrhoeae"
                     /strain="MS11"
                     /db_xref="taxon:485"
     repeat_unit     218..243
     repeat_unit     344..369
     -35_signal      373..380
     -10_signal      398..403
     RBS             455..460
     gene            465..1255
                     /gene="opaK"
     CDS             465..1255
                     /gene="opaK"
                     /note="The cloned gene is in its non-expressed state. The
                     change in the number of coding repeat unit has caused the
                     atg start codon to be out of frame with the rest of its
                     coding sequence"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="opacity protein"
                     /db_xref="PID:e145998"
     repeat_region   490..524
                     /note="signal peptide coding sequence"
                     /rpt_type=TANDEM
     mat_peptide     539..1252
                     /gene="opaK"
                     /product="opacity protein"
     terminator      1269..1294
BASE COUNT      383 a    425 c    343 g    283 t
ORIGIN

  Query Match             2.5%;  Score 81.6;  DB 1;  Length 1434;
  Best Local Similarity   95.5%;  Pred. No. 7e-08;
  Matches   84;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3117 CTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGG 3176
               ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db          1 CAGGAACGGCTTGGCACGTTACAGCTACACCGGTTCCAAACAGTACGGCAACCACAGCGG 60

Qy       3177 ACGAGTCGGCGTAGGCTACCGGTTCTGA 3204
              || | ||||||||||||||||||||||
Db         61 ACAAATCGGCGTAGGCTACCGGTTCTGA 88


RESULT 30
AE011766
LOCUS       AE011766               10807 bp    DNA     linear   BCT 29-MAY-2002
DEFINITION  Xanthomonas axonopodis pv. citri str. 306,  section 144 of 469 of
            the complete genome.
ACCESSION   AE011766 AE008923
VERSION     AE011766.1  GI:21107504
KEYWORDS    .
SOURCE      Xanthomonas axonopodis pv. citri str. 306.
```

```
  ORGANISM  Xanthomonas axonopodis pv. citri str. 306
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xanthomonas.
REFERENCE   1  (bases 1 to 10807)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Comparison of the genomes of two Xanthomonas pathogens with
            differing host specificities
  JOURNAL   Nature 417 (6887), 459-463 (2002)
  MEDLINE   22022145
   PUBMED   12024217
REFERENCE   2  (bases 1 to 10807)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-NOV-2001) Departmento de Bioquimica, Universidade de
            Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
            Brazil
FEATURES             Location/Qualifiers
     source          1. .10807
                     /organism="Xanthomonas axonopodis pv. citri str. 306"
                     /strain="306"
                     /db_xref="taxon:190486"
                     /note="pathovar: citri"
     gene            complement(149. .361)
                     /gene="XAC1343"
     CDS             complement(149. .361)
                     /gene="XAC1343"
                     /note="putative; ORF located using Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM36214.1"
                     /db_xref="GI:21107505"
                     /translation="MPLKVDGTIRHPADGDPWEYTVVVVVRNGKGEEIRRHVVDVGAM
                     HGGEERGFTLAVEASAVRSPGRRTRH"
     gene            complement(511. .1065)
                     /gene="XAC1344"
     CDS             complement(511. .1065)
                     /gene="XAC1344"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM36215.1"
                     /db_xref="GI:21107506"
                     /translation="MSMWRDQGPNKKDGVPVTPEVAAADGRLFTAEASPAPPVPAATP
                     SVVPAAAPTQRHSEAKESLIAADISIEGKIEGAGHVRLAGRFKGDVNVKGDLTIERGA
                     KLNGGVRANKVIIAGELEGNIESAAQVELQTSGVLVGDVKAGSLTVASGARMRGQADF
                     GWGDDGGKPVTTSKPVVSGDSDAT"
     gene            complement(1315. .3024)
                     /gene="XAC1345"
     CDS             complement(1315. .3024)
                     /gene="XAC1345"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="sensor histidine kinase"
                     /protein_id="AAM36216.1"
                     /db_xref="GI:21107507"
                     /translation="MSSSFTSRRRNQLALIFSALIFVVIGIGIFVGARRSLADAALVS
                     HTHEVIGRVDEIQARLLDAESAQRGFLLTGNDAYLLGYQTSVERLPILLDNLRRLIVD
                     NPAQEQHLAQLQRLVETRLRQIQHVLDIYTQSGLEPARADIRQSAFRTTSAIRELALA
                     MVQREQELLAQRAQSSEQSANLLLALALAGIPFGLVVVGMVYALLMRELRHRAQAEQL
                     AAQANRELGQSIDALQRSTADLNLLSRYTGLLQSCISAEEALMVSSRTLASLLPGVAG
                     SVYLLRASQDRAEAISHWGEPLVRSAPHLSPEQCWALRRGQPHIVQDLGRDALCAHIE
                     EPAHDASVTTACLPMSAQGTQLGFLFLSAAGPGPMPRLEIAEAAAEQLSLALSNLRLR
                     ESLRRQSIRDALTGLYNRRYLEESLSHELARCARRGLPLPVLMLDVDHFKQFNDSQGH
                     AGGDLLLAAVGELLLTRLRAEDVACRYGGEEFTVLLPEADGEEAMRVAEQIRGYIAAL
                     AVSDGARALPKVTASIGVASFPVDGEQGASLIQKADAALYRAKHAGRNRVERHGAATD
                     NAD"
     gene            complement(3135. .3461)
                     /gene="XAC1346"
     CDS             complement(3135. .3461)
                     /gene="XAC1346"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM36217.1"
                     /db_xref="GI:21107508"
                     /translation="MKTIHMLSLSAIAALTLAACQGPDAEQARRDAASAAKHAAAATE
                     QAADRAAVEAKAATADASVAARELAADASRATANAADATADKTREMANKAEQNADKHD
                     ADHAQH"
     gene            complement(3691. .4032)
                     /gene="XAC1347"
     CDS             complement(3691. .4032)
                     /gene="XAC1347"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="outer membrane protein"
                     /protein_id="AAM36218.1"
                     /db_xref="GI:21107509"
                     /translation="MTINKLLIAMALGLALTACSKPEQAQDAAASANEAAADAQTAAD
                     QAAASGSQSADAAQAAANTAAASADTAADAAQAASGAATDAAADQAADAAKTAEDTAK
                     AAKDTAEEAKK"
     gene            complement(4155. .5330)
                     /gene="atoB"
                     /note="XAC1348"
     CDS             complement(4155. .5330)
                     /gene="atoB"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="acetoacetyl-CoA thiolase"
                     /protein_id="AAM36219.1"
                     /db_xref="GI:21107510"
                     /translation="MSDIVIVAAKRTAIGSFLGQFNAVPAPTLAAAAIQGALAQSGIA
                     PADISEVIVGCVLPANLGQAPARQAAIAAGIPTSTGATTINKVCGSGMKAIMLGHDLI
                     KAGSASIVVAGGMESMSNAPHLLPNSRTGNRYGNFQAVDHMAWDGLTNPYDGQAMGVF
                     GEATAEKFGFSRADQDAFAIASVERAHAAQRSGAFADEIVPVTVATRKGDVVVDSDEQ
                     PGKSDVAKIPTLKPAFKKDGSVTAASSSSISDGAAITVLMSADDAQRRGVTPLARIVG
```

```
                     HVTHAQEPEWFTTAPVAAIQSLVSKLGWRLDDVDLFEINEAFAVVAMAPIRQLGIAHD
                     KVNVHGGACALGHPIGASGARLVVTLVNALRSRGGKRGIATLCIGGGEATAIAIELI"
     gene            5447. .8326
                     /gene="XAC1349"
     CDS             5447. .8326
                     /gene="XAC1349"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="serine protease"
                     /protein_id="AAM36220.1"
                     /db_xref="GI:21107511"
                     /translation="MGPIHGMGSGRTPMKKQDVARTVLASALAVALSACGGGGGGGGI
                     RPSTPAAPPVTSPPPPPPPTSPPPPPVTTPTTPEPAIDAHLALTNARAAQALGFTGAG
                     YRIGVIDTGINANHPALQGRVSDSFIYVDPRTNNTAVGDVVGHGTVVAELAAGRAVGQ
                     WPGGIAPGAGLVSARIISDRAPEDDGSGDGNEVDGPLGLGPVHADLISAGVRIMNNSW
                     GGLFWNDPTVTNQIAQEYRPFILSHDGLVVFASGNESRSQPSDTAALPSQPGPNGTLP
                     AADLERGWLVVGALDTANPTQLASYSNACGVAMRYCLVAPGTSLFIDPDATAGNIRYF
                     YGSGTSFAAPLVSGAAALVWQAFPYFNNDLVRQTLLGTATDLGAAGVDPVFGYGLLDV
                     GKAVLGPARFDWGTVEVDVTTLRSTWANDISGTGGLTKRGTGALVLSSTANTFTGDTQ
                     VLGGTLQTASLQSARVSIANGARLVGAGSIGGRVDNAGTLQVNGATASIAGNYTQASN
                     GRLAVNVGDRLNVAGTATIAGDLQLLGRRDYVVNNTTYAVLQATNGLQGTFDTLSSGP
                     AVTFLSATRSYDANTAYVSLQRMDVTAVAASLGAIGTASMDAAIRVEQAFQQVDAQQF
                     QGNGALSGAFIRAAAALQQSPTAKAAADSLRSLSGRAHADSAAMTFDTIDLGRRALAA
                     HFDGLSQQPRLLGNWQRALGGPGEGGATTSGFATSGWMMGNDLRLASGAVTGFAFGET
                     RSSSLGDLDGARGRDRQVQAQLYWGTTLGAAYTLGQMGFGNVNRQIERTLQLGEDRSS
                     AFSDYSGSYLSSNLETGYRLGGAHASLTPYMGLDYVRLRSDGFRESGGDGFGLRANAS
                     TSSRAQALAGVRSAYHWRGIQLGGYAEWQQALSSDGLTLDASFVGVDAWAPLRGMQPA
                     RSGGLFGITAAAPLGQQSQLRFGYDQRVGERGDDRALSLRYSADF"
     gene            complement(8388. .9488)
                     /gene="leu"
                     /note="XAC1350"
     CDS             complement(8388. .9488)
                     /gene="leu"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="leucine dehydrogenase"
                     /protein_id="AAM36221.1"
                     /db_xref="GI:21107512"
                     /translation="MLFETLDTTGHEQVIFCHNRDAGLKAIIALHSTRLGPALGGVRM
                     RPYASSEAALNDALRLSRTMTYKNALAGLNVGGGKAVIIGDPRTDKTEALFRAFGRFV
                     DTLGGRYITSEDVGTDVNDMEQIYLESEYVTGVHQVHRGSGDPAPFTAYGALQALMAS
                     LNRKLGHEEVGKASIAIQGLGHIGMELVKLLKERGAKLYVADLNQALVDRAVAEYGAE
                     AVRPDEIHQVSADVFAPCALEGAINEDTLPQLKAKIICGTANNQLASAAIGEELHRRG

  Query Match             2.4%;  Score 77.4;  DB 1;  Length 10807;
  Best Local Similarity   52.7%;  Pred. No. 7.4e-07;
  Matches  246;  Conservative    0;  Mismatches  206;  Indels   15;  Gaps    3;

Qy       1224 CGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATA 1283
               ||| || || ||  | || ||  | ||  ||||  |||| || ||| || ||||  ||
Db       6457 CGGCACCTCGTTCGCAGCGCCGCTGGTCTCCGGTGCGGCCGCGCTGGTGTGGCAGGCATT 6516

Qy       1284 CCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGG 1343
                |||    | | ||| ||       ||||     ||  |||||  ||| ||     ||| | ||
Db       6517 TCCGTATTTCAACAATGATCTGGTGCGCCAGACCTTGCTGGGAACTGCCACCGACTTGGG 6576

Qy       1344 TGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAA 1403
                || |  || || |||       ||||||| || ||||| || |  || ||||| ||
Db       6577 GGCCGCGGGTGTCGACCCGGTATTCGGCTACGGCCTGCTCGACGTCGGCAAGGCGGTGCT 6636

Qy       1404 CGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATAT 1463
                || || ||    |||     |  |||      |      | ||||    |            ||     |
Db       6637 GGGGCCGGCACGCTTCGATTGGGGCACGGTGGAGGTCGATGTCA---------CCACCTT 6687

Qy       1464 TGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAG 1523
                     |   |||       |||||||        ||||| ||||| |||| |||       |||  | |
Db       6688 GCGTTCGACCTGGGCCAACGACATCAGTGGCACCGGCGGGCTGACCAAGCGCGGCACCGG 6747

Qy       1524 CCAACTGCAACTGCACGGCA---ACAACACCTATACGGGCAAAACCATTATCGAAGGCGG 1580

                   |    ||  | ||  | |||   ||| ||||  || ||| | |||       |       |||||
Db       6748 CGCGTTGGTATTGAGCAGCACGGCCAATACCTTCACCGGCGATACCCAGGTACTCGGCGG 6807

Qy       1581 TTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCT 1640
                  |  ||  |        | ||       | |      ||       || || |||     |||| || ||||
Db       6808 CACCTTGCAGACCGCCAGCTTGCAGAGTGCGCGCGTGAGCATCG---CCAACGGCGCGCG 6864

Qy       1641 GATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCA 1687
                  |        || ||            ||| ||| |   |    ||| || ||| |
Db       6865 CCTGGTCGGCGCGGGCAGCATCGGTGGTCGGGTCGACAACGCCGGGA 6911


RESULT 31
AE012229
LOCUS       AE012229               10029 bp    DNA     linear   BCT 23-MAY-2002
DEFINITION  Xanthomonas campestris pv. campestris str. ATCC 33913,  section 137
            of 460 of the complete genome.
ACCESSION   AE012229 AE008922
VERSION     AE012229.1  GI:21112347
KEYWORDS    .
SOURCE      Xanthomonas campestris pv. campestris str. ATCC 33913.
  ORGANISM  Xanthomonas campestris pv. campestris str. ATCC 33913
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xanthomonas.
REFERENCE   1  (bases 1 to 10029)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Comparison of the genomes of two Xanthomonas pathogens with
            differing host specificities
  JOURNAL   Nature 417 (6887), 459-463 (2002)
  MEDLINE   22022145
   PUBMED   12024217
REFERENCE   2  (bases 1 to 10029)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-NOV-2001) Departmento de Bioquimica, Universidade de
            Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
            Brazil
FEATURES             Location/Qualifiers
     source          1. .10029
                     /organism="Xanthomonas campestris pv. campestris str. ATCC
                     33913"
                     /strain="ATCC 33913"
                     /db_xref="ATCC:33913"
```

```
                     /db_xref="taxon:190485"
                     /note="pathovar: campestris"
     gene            complement(103. .438)
                     /gene="XCC1295"
     CDS             complement(103. .438)
                     /gene="XCC1295"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM40593.1"
                     /db_xref="GI:21112348"
                     /translation="MNKIRMLSLSAVAVAVVTLAACQGPDAEQARRDAASAAKHAAAA
                     TEQAADRAAVQAKEATADARVAARDLAADASRATANAADATADKTREMANKAEQNAAE
                     KNAEHDAQH"
     gene            complement(690. .1031)
                     /gene="XCC1296"
     CDS             complement(690. .1031)
                     /gene="XCC1296"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="outer membrane protein"
                     /protein_id="AAM40594.1"
                     /db_xref="GI:21112349"
                     /translation="MTINKLLIAMALGLALTACSKPEQAQDAAASANEAATDAQTAAD
                     QAAASGSQSADAAQAAANTAAASADTAADAAQAASGAATDAAADQAADAAKTAEDTAE
                     AAKDTAEEAKK"
     gene            complement(1155. .2330)
                     /gene="atoB"
                     /note="XCC1297"
     CDS             complement(1155. .2330)
                     /gene="atoB"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="acetoacetyl-CoA thiolase"
                     /protein_id="AAM40595.1"
                     /db_xref="GI:21112350"
                     /translation="MSDIVIVAAKRTAIGSFLGQFNGVPAPTLAAAAIEGALAQSGVA
                     AADVSEVIVGCVLPANLGQAPARQAAIAAGIPTSTGATTINKVCGSGMKAIMFGHDLI
                     KAGSASIVVAGGMESMSNAPHLLPNSRTGNRYGNFQAVDHMAWDGLTNPYDGQAMGVF
                     GEATAEKFGFSRADQDAYAIASVERAQAAQRNGSFADEIVPVKVATRKGEIVVDSDEQ
                     PGKSDMAKIPTLKAAFKKDGTVTAASSSSISDGAAITVLMAADEAQRRGITPLARIVG
                     HVTHSQAPEWFTTAPVAAIQSLVGKIGWTLDDVDLFEINEAFAVVAMTPIKELGIAHE
                     KVNVHGGACALGHPIGASGARLVVTLVNALRSRGGKRGIATLCIGGGEATAIAIELI"
     gene            2448. .5348
                     /gene="XCC1298"
     CDS             2448. .5348
                     /gene="XCC1298"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="serine protease"
                     /protein_id="AAM40596.1"
                     /db_xref="GI:21112351"
                     /translation="MGPIHGLVRGKLKMKKTDVARTVLASALAVALTACGGGGGGGI
                     RPSTPTAPPPTSPPPPPPPTSPPPPTTAPTPPPPTTVPTPQPAFDAHLALTNTRTAQA
                     LGFTGAGYRIGVVDSGINGNHPALQGRLSDSFIYVDPRDNNIAVGDVVGHGTVVAELA
                     AGGAVGQWPGGIAPGATLVSARIIADKAPTDDGSGQGNEIDGPLGLEAVHADLIRAGT
                     RIMNNSWGGLYWTDPNVTNPVAQEYRAFVIGNDGLVVFASGNESRTEPSDTAALPSQP
                     GPNGTRPAADLERGWLVVGAVDTANPTQLASYSNACGVTMRYCLVAPGTSVYVGPDST
                     AGNLSYFYGSGTSFAVPLVSGAAALVWQAYPYFNNDLVRQTLLGTATDLGAAGVDPTF
                     GYGLLNVGKAVLGPARFDWGTVDVEVMTLRSTWANDISGNGGLTKRGSGTLVLSSTAN
                     TFIGDTQVLGGTLQTASLQSARVGVASGASLIGSGRIAGTLNNAGTVQVNGATLAVAG
                     NYIQDSTGRLALNVGDRLNVTGTASIAGELQLLGRRDYVVDNASYTLLQAAGGLQGTF
                     GTVTRGPAVTFLDASLTYDSNNAYLALRGLTATAVASTLGLSGTTAMDSAVRVEQAFQ
                     QVDVQLLQGSGAISSGFVQAAAALQQSPTGETAADSLRSLSGRAHAESAAMTFDTIDL
                     SRRALATHFDGVSQQPRLLGVWQRALGGPGEGGATSNGFATSGWMMGNDLRMASGAVA
                     GFAFGETRSNSLGDLDGARGRDRQVHAQLYWGTTRGPGYALGQMGFGNVRRQIERSLQ
                     FGLDRSAAFSDYSGSYLSGNLETGYRFGDVAASLTPYLGVDYVRLASEGFREQGGEGF
                     GLRAAASTSARTQLLAGLRSAYRWRGLQLGGYAEWQQALSSEGLSLDASFIGVDAWAP
                     LRGMQPARSGGLLGVAASAPLGERGALRFGYEQRVGERGDDRALSLKYSADF"
     gene            complement(5450. .6550)
                     /gene="leu"
                     /note="XCC1299"
     CDS             complement(5450. .6550)
                     /gene="leu"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="leucine dehydrogenase"
                     /protein_id="AAM40597.1"
                     /db_xref="GI:21112352"
                     /translation="MLFETLDTTGHEQVIFCHNRDAGLKAIIALHSTRLGPALGGVRM
                     RPYANSEAALNDALRLSRTMTYKNALAGLNVGGGKAVIIGDPKTDKSEALFRAFGRFV
                     DTLGGRYITSEDVGTDVNDMEQIYLESEYVTGVHQVHRGSGDPAPFTAYGALQALMAS
                     LNRKLGHEEVGKASIAIQGLGHIGMELVKLLKERGAKLYVADLNPALVDRAVSEYGAE
                     AVRPEEIHEVPADVFAPCALEGAINEDTLPRLKAKIICGTANNQLSSAAIGEELHKRG
                     ILYAPDYVVNAGGVMNVSLEIDGYNRERAMRLIRSIYHNLEKVFDLSTQLGIPPQQAA
                     DQIAETRIEAISKLKLPLGRTAPRFLHKLRGE"
     gene            complement(6708. .6842)
                     /gene="XCC1300"
     CDS             complement(6708. .6842)
                     /gene="XCC1300"
                     /note="putative; ORF located using Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAM40598.1"
                     /db_xref="GI:21112353"
                     /translation="MVCFDNLLADGTPDPDSLHAGLPVDAGSKWLGTLWFRQRRYRDW
                     "
     gene            complement(6983. .7483)
                     /gene="XCC1301"
     CDS             complement(6983. .7483)
                     /gene="XCC1301"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM40599.1"
                     /db_xref="GI:21112354"
                     /translation="MDTLRYLTGYPPAVLEQVRALIAQDKLGAVLAQRYPERHTVRTD
                     RQLYDYVKEMKERYMRSSPTLHKAIYDNHLQVVKHALGTHTAISRVHGSRLKASREIR
                     IAAIFRDAPAPFLQMIVAHELAHLKEPEHDKAFYQLCHHMAPDYHQLEFDLRLYLTQL
                     ALSPRD"
     gene            7692. .7883
                     /gene="XCC1302"
     CDS             7692. .7883
                     /gene="XCC1302"
                     /note="identified by sequence similarity; putative; ORF
                     located using Blastx/Glimmer/Genemark"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM40600.1"
                     /db_xref="GI:21112355"
                     /translation="MPIGDIMIKWAIIFAIIGLIAGALGFGGMAGAAMGIAKFLFWAG
                     IIIAIVLFVLGMTIAKKVT"
     gene            complement(8047. .8310)
                     /gene="XCC1303"
     CDS             complement(8047. .8310)
                     /gene="XCC1303"
                     /note="putative; ORF located using Glimmer/Genemark"

  Query Match             2.4%;  Score 77;  DB 1;  Length 10029;
  Best Local Similarity   61.0%;  Pred. No. 9.1e-07;
  Matches  125;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;
```

```
Qy    1224 CGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATA 1283
           ||| || || ||| |||  ||  | ||  ||||| | || || ||| ||  ||||    ||
Db    3479 CGGCACTTCGTTTGCCGTGCCATTGGTGTCCGGCGCCGCCGCGCTGGTGTGGCAGGCGTA 3538

Qy    1284 CCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGG 1343
           |||||   | | |||||||     ||||     ||  || ||    |||||     ||| ||||
Db    3539 CCCGTATTTCAACAACGACCTGGTGCGCCAGACCTTGTTGGGCACGGCGACCGACCTCGG 3598

Qy    1344 TGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAA 1403
            || |  ||||| ||     | |||||| |  || |||||  | | ||| |||||  ||
Db    3599 CGCCGCAGGCGTCGATCCAACGTTCGGATATGGCCTGCTCAACGTGGGCAAGGCGGTGCT 3658

Qy    1404 CGGACCCGCGTCCTTTCCGTTCGGC 1428
            || || |||    ||     |  |||
Db    3659 GGGGCCGGCGCGTTTCGACTGGGGC 3683


RESULT 32
AE003880
LOCUS       AE003880               10495 bp    DNA     linear   BCT 15-JUN-2001
DEFINITION  Xylella fastidiosa 9a5c, section 26 of 229 of the complete genome.
ACCESSION   AE003880 AE003849
VERSION     AE003880.1  GI:9105080
KEYWORDS    .
SOURCE      Xylella fastidiosa 9a5c.
  ORGANISM  Xylella fastidiosa 9a5c
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xylella.
REFERENCE   1  (bases 1 to 10495)
  AUTHORS   Simpson,A.J., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
            Alvarenga,R., Alves,L.M., Araya,J.E., Baia,G.S., Baptista,C.S.,
            Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.,
            Bueno,M.R., Camargo,A.A., Camargo,L.E., Carraro,D.M., Carrer,H.,
            Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C., Costa-Neto,C.M.,
            Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorry,H.,
            Facincani,A.P., Ferreira,A.J., Ferreira,V.C., Ferro,J.A.,
            Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R.,
            Garnier,M., Goldman,G.H., Goldman,M.H., Gomes,S.L., Gruber,A.,
            Ho,P.L., Hoheisel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P.
            and Marino,C.L.
  TITLE     The genome sequence of the plant pathogen Xylella fastidiosa. The
            Xylella fastidiosa Consortium of the Organization for Nucleotide
            Sequencing and Analysis
  JOURNAL   Nature 406 (6792), 151-157 (2000)
  MEDLINE   20365717
   PUBMED   10910347
REFERENCE   2  (bases 1 to 10495)
  AUTHORS   Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
            Alvarenga,R., Alves,L.M.C., Araya,J.E., Baia,G.S., Baptista,C.S.,
            Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.S.,
            Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M.,
            Carrer,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R.,
            Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E.,
            Docena,C., El-Dorry,H., Facincani,A.P., Ferreira,A.J.S.,
            Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C.,
            Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S.,
            Gomes,S.L., Gruber,A., Ho,P.L., Hoheisel,J.D., Junqueira,M.L.,
            Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuramae,E.E., Laigret,F.,
            Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A.,
            Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N.,
            Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L.,
            Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C.,
            Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A.,
            Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G.,
            Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C.,
            Palmieri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira
            Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V.,
            de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V.,
            Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva
            Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de
            Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M.,
            Tsuhako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S.,
            Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and
            Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP
            13083-970, Brazil
FEATURES             Location/Qualifiers
     source          1..10495
                     /organism="Xylella fastidiosa 9a5c"
                     /db_xref="taxon:160492"
                     /clone="9a5c"
     gene            312..620
                     /gene="XF0264"
     CDS             312..620
                     /gene="XF0264"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF83077.1"
                     /db_xref="GI:9105081"
                     /translation="MRELTSIEMNNVSGGDFATRLEASIVGGIAAFFAGSIWGGTRGG
                     DGGGVLGIGSIGQGVGMVYGGIAGAIGGAIGCFVLDKDVTYSYTTGFMNSLFSGNFAK
                     "
     gene            732..953
                     /gene="XF0265"
     CDS             732..953
                     /gene="XF0265"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF83078.1"
                     /db_xref="GI:9105082"
                     /translation="MLIANTFYFIAVYCPQPSMIQKTFGCRIPAGNNDVLMSLLKHVI
                     RMEFAACLGGAFQKKHYYYACCLLFICNA"
     gene            complement(2562..2858)
                     /gene="XF0266"
     CDS             complement(2562..2858)
                     /gene="XF0266"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF83079.1"
                     /db_xref="GI:9105083"
                     /translation="MVSVSVSVSALVSALVSALALASVSVSVSVSVSVSVSVSVSVSV
                     SVSVSVSVLALALVSMGELTHYHHRHSRKKLTLISGCYWEKMSFSFLSFNEDNQ"
     gene            2858..5590
                     /gene="XF0267"
     CDS             2858..5590
                     /gene="XF0267"
                     /note="similar to GI|4115629 (percent identity: 28 %/query
                     alignment coverage:  99.3 %/subject alignment coverage:
                     87.3 %); identified by sequence similarity; putative; ORF
                     located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="serine protease"
                     /protein_id="AAF83080.1"
                     /db_xref="GI:9105084"
                     /translation="MPALDPALAAKFRVHLDGTNTEPAHTLGFTGKGYMIGVVDSGVN
                     RNHVTLHDKVVHNNVYVDYQDPASPDDVSGHGTNVAQLIAGQPVEYWPGGIAPGATIF
                     SARVFSSAAENAKQGPIGGNFFEGEAAQLKRVNADMIAAGVRIQNNSWGYENQNTGDK
                     QVWIDPAVTDAFVNVYRDFVLNNNGLVIFASGNYSQQQPGRLSRLPSLPTVEGGASPA
                     DLERGWLVVAALDSFDHPDQLTSYSNHCGIAMHYCLAAPGDLIEIDPNVTTSPVDGDK
                     GYLNGYLIQNGTSFAAPLVSGAAALVWEAFPYFSNDLVRQTLLGTATDLGAPGVDEVF
                     GYGALNVGKAVLGPAKFDWGDVQVSFDDRRSTWGNDITGSGGLIKRGTGTLVLGGSNN
                     QYTGATQVLGGTLQASSLGASAVSVSNAATLIGSGRFGGAVSNAGTLELGKDGLKVQG
                     DYTQLETGRLALHVGDQLSVAGNATLKGGELQVLGKRDYVTFNTSYSVLQADGSLTGT
                     FSQLTWGPAVLLGEGRLTYGANNAYVTLQRLNVSAAAQALGIVDRVAQASAQRVEQAF
                     RAIDAQQAQGRGALPTSFIGAAAALQQSSSAAVAAGSLQSLSGQSHVAAALVSLDAID
```

```
                 LGRRSLATRLDAPRSGERIVTWHQALGGPGQNGAASGTFSLSGWLVGHDAQLASGDIV
                 GVAFGQTDTSPSGSAFGSRGIDRQVQGRVYLQRTQGPLYVLGQLGFGSFQRRLDRQLQ
                 LGDWNDWTSSRYSGQFWSGSVEAGYHWHQGTLALTPYLGLDQTQLRTDGFREQSGSGF
                 GLQVQSAQATRSQLLAGVRTEWRWAGVQWRGYGEWQQTLRQSGLNPQASFTATSSWTP
                 LVASSWPGRSGGVLGLSMVLPVGVTAGLWTMGLEHYFGTRSRGESLGLRYQLGF"
     gene        complement(6065. .7252)
                 /gene="XF0268"
     CDS         complement(6065. .7252)
                 /gene="XF0268"
                 /note="similar to SP|P39381 (percent identity: 56 %/query
                 alignment coverage:  95.9 %/subject alignment coverage:
                 96.7 %); identified by sequence similarity; putative; ORF
                 located using Glimmer/RBSfinder"
                 /codon_start=1
                 /transl_table=11
                 /product="conserved hypothetical protein"
                 /protein_id="AAF83081.1"
                 /db_xref="GI:9105085"
                 /translation="MKTLEEKKSASNLNHLKTALFGMIILSLGIGIGRFLHTPILPVM
                 LREGQFTFSQLSYIASANYAGYLLGSLFFSFGKLGNTSRTTMMLYGAAIVTNVLIFAM
                 AFTSHFFLVMLIRFAAGVSSAAMMIFGSITIMRHTFNVRVIASLYAGVGIGILLGNEY
                 VVIGLRHGLNASALWYGASVLSFILLLLLFVLTPHVEDKPREASASFPVQQNSFLWWQ
                 LAALYGCAGFGYIIIATYLPLIAKTFNVPFIAEHLWSLVGLTIIPSCFAWLWAAQRWG
                 TRRCLTINLLIQGCCVLLTLLSQAPFLLVISCIGFGATFMGTTSLVMPLTRQVSAPHR
                 INLLGLVTLTYGIGQILGPLLTSLLHSRLNTVTPAIICGAVALFVAAGICQYCLDKKA
                 VNV"
     gene        7299. .7409
                 /gene="XF0269"
     CDS         7299. .7409
                 /gene="XF0269"
                 /note="hypothetical protein; identified by sequence
                 similarity; putative; ORF located using Glimmer/RBSfinder"
                 /codon_start=1
                 /transl_table=11
                 /product="hypothetical protein"
                 /protein_id="AAF83082.1"
                 /db_xref="GI:9105086"
                 /translation="MGKPSGPCVAEISVLLGHGSDRVCGIRIDVTVQCLE"
     gene        complement(7542. .7700)
                 /gene="XF0270"
     CDS         complement(7542. .7700)
                 /gene="XF0270"
                 /note="hypothetical protein; identified by sequence
                 similarity; putative; ORF located using Glimmer/RBSfinder"
                 /codon_start=1
                 /transl_table=11
                 /product="hypothetical protein"
                 /protein_id="AAF83083.1"
                 /db_xref="GI:9105087"
                 /translation="MHHATQAQQYDNLNTLQFKYQWIDVTSCINTGYAHVQRADQSMH
                 ALNTASQT"
     gene        7925. .8020
                 /gene="XF0271"
     CDS         7925. .8020
                 /gene="XF0271"
                 /note="hypothetical protein; identified by sequence
                 similarity; putative; ORF located using Glimmer/RBSfinder"
                 /codon_start=1
                 /transl_table=11
                 /product="hypothetical protein"
                 /protein_id="AAF83084.1"
                 /db_xref="GI:9105088"
                 /translation="MNYFVIQNGRFIERYRCEVFSMSLALLFSLF"
     gene        complement(8158. .8733)
                 /gene="XF0272"
     CDS         complement(8158. .8733)
                 /gene="XF0272"
                 /note="hypothetical protein; identified by sequence
                 similarity; putative; ORF located using Glimmer/RBSfinder"
                 /codon_start=1
                 /transl_table=11
                 /product="hypothetical protein"
                 /protein_id="AAF83085.1"
```

```
  Query Match             2.3%;  Score 72.8;  DB 1;  Length 10495;
  Best Local Similarity   51.6%;  Pred. No. 8.5e-06;
  Matches  223;  Conservative    0;  Mismatches  197;  Indels   12;  Gaps    2;

Qy       1224 CGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATA 1283
              ||| || || ||| | || ||  | ||  | ||  | |||||  || |   | ||  ||
Db       3715 CGGCACCTCGTTTGCAGCGCCGTTGGTTTCTGGTGCAGCGGCATTGGTATGGGAGGCATT 3774

Qy       1284 CCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGG 1343
               || |   | |||||||||    ||||    || |||||    |||||    ||  | ||
Db       3775 TCCCTATTTCAGCAACGACCTGGTGCGGCAGACCCTGCTTGGCACGGCGACTGATCTTGG 3834

Qy       1344 TGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAA 1403
               ||    |||||||||        ||||||| ||     ||| ||| |||||| ||   ||
Db       3835 CGCCCCCGGCGTGGATGAAGTCTTCGGCTATGGTGCACTGAATGTGGGTAAAGCAGTGCT 3894

Qy       1404 CGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATAT 1463
               || || ||     ||     |   | ||  |     | |            |     || |
Db       3895 GGGCCCGGCTAAGTTCGATTGGGGTGATGTCCAAGTC---------AGCTTCGATGACAG 3945

Qy       1464 TGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAG 1523
                  | | |||    | || |||||  | ||||  |||||| |||||||     ||    | |
Db       3946 GCGTTCCACCTGGGGCAATGACATCACCGGCAGCGGCGGCTTGATCAAGCGTGGTACCGG 4005

Qy       1524 CCAACTGCAACTGCACGGCAACAACA---CCTATACGGGCAAAACCATTATCGAAGGCGG 1580
                 ||||   ||     |||| ||| |          || || ||     |||      ||    || ||
Db       4006 AACACTGGTGCTTGGTGGCAGCAATAATCAATACACCGGAGCCACCCAAGTCCTGGGAGG 4065

Qy       1581 TTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAAGGTGCGCT 1640
              | |  ||  |  || | ||          ||| |    |  ||||     | |    |   || |
Db       4066 TACATTGCAGGCGTCCAGCTTGGGTGCATCCGCAGTCAGCGTTTCTAACGCTGCGACGTT 4125

Qy       1641 GATTTATAACGG 1652
              ||||  || |||
Db       4126 GATTGGTAGCGG 4137


RESULT 33
AB015053
LOCUS       AB015053               14988 bp    DNA     linear   BCT 12-NOV-1999
DEFINITION  Pseudomonas fluorescens genes for ABC exporter operon, complete
            cds.
ACCESSION   AB015053
VERSION     AB015053.1  GI:4115622
KEYWORDS    ORFX; extracellular lipase; serine protease homologue; outer
            membrane protein (OMP); Membrane fusion protein (MFP); ABC protein;
            protease inhibitor; alkaline protease.
SOURCE      Pseudomonas fluorescens (isolate:No.33) DNA.
  ORGANISM  Pseudomonas fluorescens
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (sites)
  AUTHORS   Kawai,E., Idei,A., Kumura,H., Shimazaki,K., Akatsuka,H. and
            Omori,K.
  TITLE     The ABC-exporter genes involved in the lipase secretion are
            clustered with the genes for lipase, alkaline protease, and serine
            protease homologues in Pseudomonas fluorescens no. 33
  JOURNAL   Biochim. Biophys. Acta 1446 (3), 377-382 (1999)
  MEDLINE   99453738
REFERENCE   2  (bases 1 to 14988)
  AUTHORS   Kumura,H. and Omori,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-1998) Kenji Omori, TANABE Seiyaku Co. Ltd,
            Discovery Research Laboratory; 2-50 Kawagishi-2-chome, Toda,
            Saitama 335-8505, Japan (E-mail:k-omori@tanabe.co.jp,
            Tel:81-48-433-8041, Fax:81-48-433-8157)
FEATURES             Location/Qualifiers
     source          1. .14988
                     /organism="Pseudomonas fluorescens"
                     /isolate="No.33"
                     /db_xref="taxon:294"
     -35_signal      35. .40
```

```
-10_signal      59..64
-35_signal      164..169
-10_signal      187..192
RBS             243..249
gene            257..1705
                /gene="aprAPF33"
CDS             257..1705
                /gene="aprAPF33"
                /codon_start=1
                /transl_table=11
                /product="alkaline protease"
                /protein_id="BAA36461.1"
                /db_xref="GI:4115623"
                /translation="MSKVKDKAIVSAAQASTAYTQIDSFSHQYDRGGNLTVNGKPSFT
                VDQAADHLLRENAAYRDVDGNGKIDLTYTFLTSASNATMNKHGITGFSQFNTQQKAQA
                VLAMQSWADVANVTFTEKASGGDFHMTFGNYSGGQDGAAAFAYLPGTNAKYNESGLDG
                TSWYLTNSSYTPNKTPDLNNYGRQTLTHEIGHTLGLDHPGDYNAGTGNPSYKDADYGQ
                DTRGYSVMSYWSESNTNQNFSKGGVEAYSSGPLIDDIAAIQKLYGANYNTRAGDTTYG
                FNSNTGRDFLSATSSADKLVFSVWDGGGNDTLDFSGFTQNQKINLNETSFSDVGGLVG
                NVSIAKGVTVENAFGGSGNDLLIGNAAANILKGGAGNDIIYGGGGADQLWGGAGSDTF
                VFGSSSDSKPGAADKIFDFTSGSDKIDLTGITKGAGLTFVNAFTGHAGDAVLSYASGT
                NLGTLAVDFSGHGVADFLVTTVGQAAVSDIVA"
RBS             1765..1767
gene            1774..2148
                /gene="inhPF33"
CDS             1774..2148
                /gene="inhPF33"
                /note="putative"
                /codon_start=1
                /transl_table=11
                /product="protease inhibitor"
                /protein_id="BAA36462.1"
                /db_xref="GI:4115624"
                /translation="MVMTLFRTTVAWCFQAMLMSAGASAMASSLVLPTSAQLAGHWAL
                HQQEQVCALDLIEQANALGGDVACAGQWLGEAPVSWTPTPDGIWLMNAEGTGITHLNR
                QKEGEYEGRSKSGAVLLLQRAP"
RBS             2236..2243
gene            2253..4004
                /gene="aprDPF33"
CDS             2253..4004
                /gene="aprDPF33"
                /function="ABC exporter"
                /codon_start=1
                /transl_table=11
                /product="ABC protein"
                /protein_id="BAA36463.1"
                /db_xref="GI:4115625"
                /translation="MAKTTPVAPLFKALGDYKSILISIGCFTALINVLMLVPSIYMLQ
                VYDRVLSSQNETTLVMLTLMVVGFFAFIGVLEVIRSFIVIRIGSQLERRFNLRVYKAA
                FERNLQRGEGHGGQSLGDLTHIRQFITGPALFAFFDAPWFPIYLFVIFLFNVWLGVLA
                SAGAVLLIGLACLNEYLTKKPLGEASGYSQQSTQLATSHLHNAETIQAMGMLGALRKR
                WFKVHSQFLGLQNKASDTGSIISSLSKSLRLCLQSLVLGLGALLVIKGDMTAGMMIAG
                SILMGRVLSPIDQLIAVWKQWSSAKLAYQRLDNLMREFPPQGEQMALPAPKGNVSFEQ
                VSAGPPGRRVPTLHQVSFNLAAGEVLGVLGASGSGKSTLARVLVGVWPTLAGTVRLDG
                ADIHRWDRDDLGPHIGYLPQDIELFSGSIADNIARFRDADPQLVVQAAQQAGVHELIL
                RLPHGYDTVLGDEGSGLSGGQKQRVALARALYGGPRLIVLDEPNSNLDTVGEAALASA
                IMQMKAQGSTVVLVTHRSSALAQADKLLVLNEGRLQAFGPSQDVLRALSGQQEAPREK
                AGVSLSRQYQAGRNPGA"
RBS             3986..3991
                /gene="aprDPF33"
gene            4001..5335
                /gene="aprEPF33"
CDS             4001..5335
                /gene="aprEPF33"
                /function="ABC exporter (MFP)"
                /codon_start=1
                /transl_table=11
                /product="Membrane fusion protein (MFP)"
                /protein_id="BAA36464.1"
                /db_xref="GI:4115626"
                /translation="MSSDSSIQINRDYEQTHPEHGARFFARMGWLLTVVGAGGFFLWA
                SLAPLDQGIPVQGTVVVSGKRKAVQTLSPGVVSRILVREGEAVKQGQPLFRLDQTQNQ
                ADVHSLQAQYRMAWASVARWQSERDNQSTITFPAELSGNPDQALALVLEGQRQLFSSR
                REAFAREQAGIRANIEGATAQLNGMRRARSDLTAQAQSLRDQLNNLQPLADNGYIPRN
                RLMEYQRQLSQVQQDLAQNTGESGRVEQGILESRLKLQQHSEEYQKEVRSQLADAQLR
                SLTLEQQLTSAGFDLQHSEINAPADGIAVNLSVHTEGAVVRAGETLLEIVPQDTRLEV
                EGRLPVHLVDKVGTHLPVDILFTAFNQSRTPRVPGEVSLISADQMLDEKTGMPYYVLR
                TTVSSSALEKLHGLVIKPGMPAEMFIRTGERSLLNYLFKPLLDRAGSALTEE"
RBS             5327..5332
                /gene="aprEPF33"
gene            5338..6672
                /gene="aprFPF33"
CDS             5338..6672
                /gene="aprFPF33"
                /function="ABC exporter (OMP)"
                /codon_start=1
                /transl_table=11
                /product="outer membrane protein (OMP)"
                /protein_id="BAA36465.1"
                /db_xref="GI:4115627"
                /translation="MKPMFIALLLACSSAQAAMGPFDVYEQALRNDPVFLGAIKERDA
                GLENRTIGRAGLLPKLSYNYNKGRNNSQATLPDGRGGNYHDDRNYNSYGSTFTLQQPL
                FDYEAYANYRKGVAQALFADESFRDKSQALLVRVLTYYTQALFAQDQIDIARAKKKAY
                EQQFQQNQHLFQQGEGTRTDILEAESRYELATAEEIQALDEQDASLRELGALIGVQSV
                NINDLAPLNPGFAAFTLTPANYDTWHELALTNNPTLASQRQSLEVARYEVERNRAGHM
                PRVTAYASSRQQESDSGNTYNQRYDTNTIGVEISMPLYAGGGVSASTRQASRAMEQAE
                YELEGKTRETLIELRRQFSACLSGVSKLRAYQKALTSAEALVVSTKQSILGGERVNLD
                ALNAEQQLYSTRRDLAQARYDYLMAWTKLHYYAGNLRDTDLAKVDEAFGPEH"
RBS             6692..6697
gene            6706..9663
                /gene="pspA"
CDS             6706..9663
                /gene="pspA"
                /codon_start=1
                /transl_table=11
                /product="serine protease homologue"
                /protein_id="BAA36466.1"
                /db_xref="GI:4115628"
                /translation="MITDSPRFKPFTAGSLLLLSVAAHGQYLEAGKPGDPASWRSTEF
                QSDWGLDRMKANEAYAAGITGSGVKIGALDSGFDPNHPEASKDRYHAVTASGTYVDGS
                PFNTTGALNPNNDSHGTHVTGTMGAARDGVGMHGVAYNAQIYVGNTNANDSFLFGPTP
                DPQYFKAVYSALVDSGVRAINNSWGSQPKDVSYQTLGDLHAAYAQHYNQNTWLDAAAD
                VAKAGVINVFSAGNSGYANASVRSALPYFQPGLEGHWLAVSGLDKANNQKYNQCGIAK
                YWCISTPGALINSTIPEGGYGVKSGTSMSAPHATGALALVMERYPYLNNEQALQVLLT
                TATQLNGAVTDAPTTQVGWGVPDLGRAMHGPGQLLGPMTVTLAAGQGDVWSNAISDKA
                LIQRQAEDRAETLAWQQTLRDKGWENGVGAGASQQDQTDYAVGTARAAAAAQRIYEGS
                LTKSGAGSLILTGDNTYRGPTTVNGGRLAVNGSLTSAVTVNDSGTLGGSGRIAALQVN
                NGGTVAPGNSVGTLQVAGDVNLAAGSTYAVELTPTRSDQIVATGKATLGGGTVTLALE
                NSPTLLSQQQAESLVGRQYTILQAAGGIQGQFGAVLPNYLFLGGTLDYAANGVQLAIA
                RSDASFASVGATRNQRSVAAAAEQLGAGNPVYESILRSDSLATAQQGFQQLSGEIYPA
                IGTMLINDSRQVRDAVGERLRHTPVAGESNVWVKALGSWGKTDSRSETAGYTSSLGGM
                LAGVDGALDDQTRVGLVAGYSDSSLNMGSGTHSSASIDSYHLGAYAGHELGNWRLSVG
                GAYSWHRGDVKRDLQYGEVSGKQKARLDARSAQVFTEAAYRLNVPAVALEPFANLAYV
                HLDNERFHEKGDAAALARGSDSRDAVLSTLGLRASKTLALNDRQQLDVSGSLGWQHSL
                SAVESEEHLAFVAGGPSFAVQSSPLMRDAALVGVQASLALSKSTRVNLDYTGQLAGRE
                KIQGVGLSLNWQF"
RBS             9835..9842
gene            9848..12958
                /gene="pspB"
CDS             9848..12958
                /gene="pspB"
                /codon_start=1
                /transl_table=11
                /product="serine protease homologue"
                /protein_id="BAA36467.1"
                /db_xref="GI:4115629"
                /translation="MKTTKRGFQSAIPTGPGYPLKALSGVPYGALLCCLASLGTAQAA
                PYVETGKLGDAASWRSNEFKADWGLGAVHADAAYAAGYTGKGVKLGIFDQPVYAQHPE
                FASPGKVVTIVTEGIRQYTDPYIPVKAGDAFRYDGTPSLGSNGKLGNHGTHVGGIAAG
                NRDGGPMHGVAFDAQIISAENGDPGPEDGIILGNDGAVYKAGWDALVASGARIINNSW
                GIGIGDQYAKGGRDPAFPNFTVNEAQAQFNNIRPILGTIAGGAYQGAIDAARSGVLTI
                FAAGNDYNRNNPDAISGLAYFVPEIAPNWLSVAALQQNPDTASADPYVISTFSSRCGY
                AASFCVAAPGTKIYSSVINGTTLENLTTDYANFNGTSMAAPHVAGSAAVLMERFPYMS
                GDQISTLLKTTATDLGAPGIDSLYGWGMINLGKAINGPGMFVTAEDIPAEFRIDGAYG
                SGQFVADLPGMGAVVDAGKPTQRLCTDVHCGLDVWSNDIAGHGGLTKQGIGALVLTGT
                NTYSGPTLVNQGLLAINGSITSDVTVGSSGVVGGSGHIGSLYANSGGTVAPGNSIGTL
```

```
                     NVAGDVSFAAGSTYAVELSATNSDRIVAGGKAILSGGTVTLALENSPTLLTQTQAESL
                     IGRQYNILQAAGGITGSFGSVLPNYLFIGGTLNYAANGVQLDVARNANTFASVAVTPN
                     QRSVAAAAEQLGAGNAVYESLLSAPNAASAQGAFQQLSGEIYPALETALVNDNRYLRE
                     AVGERLQQGEMGASTETIDSRGNVWVKALGAWGKTDSRSDTAGYTTSIGGMLAGVDGA
                     LDEDTRLGLVAGYSDTSLNMGSDTHSRASVDSYHLGAYAGHEIGALRLSGGATYSWHR
                     ADVKRELQYGEVAGKQKAKVDARSTQVFTEAAYRLNLQPLALEPFANLAYVHLDADGF
                     TEKGDAAALKGGDDSRDLVLSTLGVRALKTLNVSDHQQLELSGTLGWQHNLSSIDSER
                     HLAFASGSPSFSVESSPMVRDAALVGARVSLALSKDAKVNLDYNGLLASKEKVHGVGL
                     SLDWAF"
     RBS             12993. .13002
     gene            13010. .14440
                     /gene="lipAPF33"
     CDS             13010. .14440
                     /gene="lipAPF33"
                     /codon_start=1
                     /transl_table=11
                     /product="extracellular lipase"
                     /protein_id="BAA36468.1"
                     /db_xref="GI:4115630"
                     /translation="MGVFDYKNLGTEGSKALFADAMAITLYSYHNLDNGFAVGYQHNG
                     FGLGLPATLVGALLGSTDSQGVIPGIPWNPDSEKAALDAVHKAGWTPISASTLGYGGK

  Query Match             2.3%;  Score 72.8;  DB 1;  Length 14988;
  Best Local Similarity   59.0%;  Pred. No. 8.7e-06;
  Matches  125;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy       1224 CGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATA 1283
              ||| || ||| |  | || |||   ||  |||| |  |||||  || || || ||   | 
Db      10951 CGGCACCTCCATGGCTGCGCCCCATGTGGCCGGGAGTGCGGCGGTGTTGATGGAGCGCTT 11010

Qy       1284 CCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGG 1343
              |||||  ||||||  |||| |  |    ||   |||   ||| || ||    ||| ||||
Db      11011 CCCGTACATGAGCGGCGACCAGATTTCCACGCTGCTCAAGACCACCGCGACCGACCTCGG 11070

Qy       1344 TGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAA 1403
               ||    ||| | |||     || |||||||||  || |  |   ||| |||||||| ||
Db      11071 CGCGCCGGGCATCGACTCGCTGTACGGCTGGGGCATGATCAACCTGGGCAAGGCCATCAA 11130

Qy       1404 CGGACCCGCGTCCTTTCCGTTCGGCGACTTTA 1435
              ||| || |  |  |||  |   |  ||   ||
Db      11131 CGGCCCGGGGATGTTTGTGACGGCTGAGGATA 11162


RESULT 34
AF286062
LOCUS       AF286062                12152 bp    mRNA    linear   BCT 07-MAR-2001
DEFINITION  Pseudomonas brassicacearum phase variable exoenzymes operon,
            complete sequence.
ACCESSION   AF286062
VERSION     AF286062.1  GI:9438186
KEYWORDS    .
SOURCE      Pseudomonas brassicacearum.
  ORGANISM  Pseudomonas brassicacearum
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 12152)
  AUTHORS   Chabeaud,P., de Groot,A., Bitter,W., Tommassen,J., Heulin,T. and
            Achouak,W.
  TITLE     Phase-variable expression of an operon encoding extracellular
            alkaline protease, a serine protease homolog, and lipase in
            Pseudomonas brassicacearum
  JOURNAL   J. Bacteriol. 183 (6), 2117-2120 (2001)
  MEDLINE   21125577
   PUBMED   11222613
REFERENCE   2  (bases 1 to 12152)
  AUTHORS   Chabeaud,P., Heulin,T. and Achouak,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) DSV/DEVM/LEMiR, CNRS-CEA UMR 163, CEA
            Cadarache, Saint-Paul-Lez-Durance 13108, France
FEATURES             Location/Qualifiers
     source          1. .12152
                     /organism="Pseudomonas brassicacearum"
                     /strain="NFM421"
                     /db_xref="taxon:86264"
     gene            257. .12051
                     /gene="phase variable exoenzymes operon"
     gene            257. .1720
                     /gene="aprA"
     CDS             257. .1720
                     /gene="aprA"
                     /codon_start=1
                     /transl_table=11
                     /product="AprA"
                     /protein_id="AAF87588.1"
                     /db_xref="GI:9438187"
                     /translation="MFMSKVKSSAIDSAEQLLWAPQSLAAASSAFNQINSFSHQYDRG
                     GNLTVNGKPSYSVDEAATQLLRDGAAYQDRDGSGKIELTYTFLTSASSSTMNKHGISG
                     FSQFSAQQKGQAALAMQSWADVANVTFAEKASGGDGHMTFGNYSGGQDGAAAFAYLPG
                     TGAGYDGTSWYLINSSYTQNKNPDLNNYGRQTLTHEIGHTLGLAHPGDYNAGEGSPTY
                     DDATYGQDTRGYSVMSYWSESNTDQNFSKGGVEAYASGPLMDDIAAIQKLYGANTTTR
                     TGDSTYGFNSNTGRDFLSASSSSDKVVFSVWDAGGRDTLDFSGFTQNQKINLNDASFS
                     DVGGMVGNVSIAKGAVIENAIGGAGSDLLIGNSVANELKGGAGNDILWGAGGADKLWG
                     GSGSDTFVFAASSDSRPGATDQILDFVSGLDKIDLTGITKGAGLHFVSALTGAAGDAV
                     LSTSAGNSLLSVDFSGHGVADFLVSTVGQAAVSDIVA"
     gene            1792. .2214
                     /gene="aprI"
     CDS             1792. .2214
                     /gene="aprI"
                     /codon_start=1
                     /transl_table=11
                     /product="AprI"
                     /protein_id="AAF87589.1"
                     /db_xref="GI:9438188"
                     /translation="MPSSVQATAGLLATLMMFCGEVAMARSLLLAEPSQLAGQWQAVL
                     SSPQDNAQTQAMQDKPSNSCLVELKVDQTLGGQTDCLGQWLGDEPVRWFTEPDGLSLI
                     GKQDSRTHLGLRQGDHYQMTLKSGLILRLERNKSQSAH"
     gene            2354. .4153
                     /gene="aprD"
     CDS             2354. .4153
                     /gene="aprD"
                     /codon_start=1
                     /transl_table=11
                     /product="AprD"
                     /protein_id="AAF87590.1"
                     /db_xref="GI:9438189"
                     /translation="MNRVRSVAAAPLFKALGDYKNILISVGCFTALINVLMLAPSIYM
                     LQVYDRGLSSQNETTLAMLSLMVVGFFVFIGLLEMVRSFVVIRIGSQLERRFNLQVYK
                     AAFERNLRQGEGNAGQSLGDLTHIRQFLTGPALFAFFDAPWFPIYLLVIYLFNVWLGV
                     FASIGTLLLIGLACLNEAMTKRALGQASVYSQQSTQLATSHLHNAETIQAMGMLGALR
                     SRWFTVHSRFLGLQNQASDTGAVISSLSKTLRLCLQSLVLGLGALLVIKGEMTAGMMI
                     AGSILMGRVLSPIDQLIAVWKQWSGAKLAYRRLDALLQAFAPQDDGMTLPAPKGQVSF
                     EQVSAGPPGQRNATLQQVSFSLAAGEVLGVLGASGSGKSTLARVLVGVWPTFAGTVRL
                     DGADIHRWNREQLGPHIGYLPQDIELFSGSIAENISRFRQADPQKVVAAAQQAGVHEL
                     ILRLPQGYDTVLGEDGGGLSGGQKQRVALARAMYDRPSLVVLDEPNSNLDTVGEAALA
                     SAIAQMKTQGTSVVLVIHRSSALAQADKLLVLDEGRLQAFGPSQQVLRALSGQSGQQD
                     NAPRSSPPNSPQTAPNGLSMSRQYQAASKPGGT"
     gene            4168. .5484
                     /gene="aprE"
     CDS             4168. .5484
                     /gene="aprE"
                     /codon_start=1
                     /transl_table=11
                     /product="AprE"
                     /protein_id="AAF87591.1"
                     /db_xref="GI:9438190"
                     /translation="MNLEHSQDVEHPERDARFFTRMGWALALVGAGSFFTWASLAPLD
                     QGIPVQGTVVVSGKRKAVQSMSSGVVSRILVREGEVVKQGQPLFRLDQTQAAADVQSL
                     QAQYRMTWASLARWQSERDNLLQVSFPAELSRDPDPRLALVLEGQRQLFSSRREAFAR
                     EQAALRANIEGAAAQLAGMRRARSDLTAQAESLRLQLGNLQPLADNGYIPRNRLLDYQ
                     RQLSQVQQELAQNSGESGRVEQGIVESRLKLQQHHEEYQKEVRSQLADAQLKSETLQQ
                     QLKSAGFELQHSEILATADGIAVNLGVHTEGAVVRQGDTLLEIVPQGTRLEVEGHLPV
                     HLIDKVGPHLPVDILFTAFNQNRTPRVPGEVSLISADQMLDEKTGAPYYVLRSSVSDV
                     AMDKLNGLVIKPGMPAEMFVRTGERSLLNYLFKPLLDRAGSALTEE"
     gene            5487. .6848
                     /gene="aprF"
     CDS             5487. .6848
```

```
                     /gene="aprF"
                     /codon_start=1
                     /transl_table=11
                     /product="AprF"
                     /protein_id="AAF87592.1"
                     /db_xref="GI:9438191"
                     /translation="MSGGMKRFYWLAVLAVSVCNNAWAMGPFEFYEQALRNDPVYLGA
                     IKERDAGLENRAIGRAGLLPHLGYSYNKGRNQSKVTYLNDRGASQHDDRNYSSYGSSL
                     TLQQPLLDYEAYAAYRKGVAQALFADESFRGKSQELLVRVLSYYTQALFAQDQIDIAL
                     AKKKAFEQQFHQNEHLFRQGEGTRTDILEAQARYELAIAEEIEARDEQDAALRELGSL
                     IGVPALDIGDLAPLHDTFQTFALQPASFDTWHELAVSNNPNLASQRQAVDVARFEVER
                     NRAGHLPKVSAYATMRQNESESGNTYNQRYDTNTIGLEVSVPLYAGGGVSASTRQASR
                     NMEQAEYELDAKTRETLIELRRQFSACVSGANKLRAYQKALSSAEALVVSTRQSILGG
                     ERVNLDALNAEQQLYTTRRDLAQARYDYLMAWTKLHYYAGTLGPQDLAKVDEAFVSRA
                     PIQ"
     gene            7036..10128
                     /gene="pspB"
     CDS             7036..10128
                     /gene="pspB"
                     /codon_start=1
                     /transl_table=11
                     /product="PspB homolog"
                     /protein_id="AAF87593.1"
                     /db_xref="GI:9438192"
                     /translation="MKVDNKKQPGQTGRVFLLKALNRAVLCALATLGTAQAAPYVESG
                     RAGNPDSWRSSEFNAEWGLGAINAQQAYAAGYTGKGVKLGIFDQPVYAAHPEFSGTGK
                     VVTLVTSGIREYTDPYIPVKAGDAFRYGGSPTVGSDGKLGSHGTHVGGIAAGSRDGSP
                     MRGVAYNAQIISADNGDPGPEDGIILGNDGAVYKAGWDALIASGARIINNSWGIGITD
                     RFDLGGRDPAFPHFTVQDAQLQFDQIKPLLGTKPGGAYEGAIAAARSGIVTIFAAGND
                     YNLNNPDAIAGLAYFVPDIAPNWLTVAALQINPDTASADPYTISTFSSRCGYTASFCV
                     SAPGTRVYSAVIGGTSADDLTVGYANKSGTSMAAPHVAGSVAVLMERFPYMTGAQVAS
                     VLRTTATDMGAPGVDALYGWGMINLGKAIDGPSMLVTEQDIPEEFRIEGAYGSGQFVV
                     DLPGIGAIIDAGKSTERLCSGIQCGLDVWRNDIAGHGGLTKEGIGTLVLTGANTYSGP
                     TLVNQGRLAVNGSLASAVTVNDGGILGGNGRIASLTANRGGSVAPGNSIGTLQVAGDV
                     TFQPGSNYAVEVSPTSSDLIVAGGKASIDGATVSLSLENSPTLLTAGDVKACWAGSSI
                     SCRPSMVLTGASHEVLPNYAFLGGDLAYSASGIQLAVERNDASFASLGLTPNQRSVAS
                     AAEQLGAGNGLYEALLLSPSAAVAQQAFQQLSGEIHPAIGTLLINDSRQLRDAVGDRL
                     RQDALYDAGTPTDASSNAWFKVLGAWGKNDGGHDHASTTSSIGGLLAGVDGLISDQTR
                     LGFVTGYSDSSLSMGDGTHSSAKADSYHLGVYLGHEIDALRLTVGGAHSWHRIDVKRD
                     LQWGDVSGREKTKRDARTTQLFTEAAYRLDLQPVALEPFANLAYVHLDSDSFHEKGDN
                     AALHAGDDRRDAVLSTLGVRASHTLALSDKQQLQLSGSLGWQHNLSSTTSEQDLAFAD
                     SSSTFAVQSVSMDRNAAVIGARAGLAVAKDVRVNLDYNGLLGSRDKSHGVGLTLDWQF
                     "
     gene            10192..12051
                     /gene="lipA"
     CDS             10192..12051
                     /gene="lipA"
                     /codon_start=1
                     /transl_table=11
                     /product="LipA"
                     /protein_id="AAF87594.1"
                     /db_xref="GI:9438193"
                     /translation="MGIYDYKNLGTTESKALVSDAMAIMLYSYHNLDNGFAAGYQHNG
                     FGAGLPATLVTALLGGTDSQGVIPGVPWNPDSEKLALEAVQKAGWTPISADQLGYAGK
                     VDGRGTFFGEKAGYTTAQVEILGKYDAEGHLQEIGVSFRGTSGPRENLIGDSIGDVIN
                     DLMAALGPKDYRKNYVGEAFGNLLGDVAKFAQAHGLTGNDVLVSGHSLGGLAVNSLAD
                     LSASKWSGFYSSANYIAYASPTQSSTDKVLNIGYENDPVFRALDGSTFNLASVGVHDG
                     QQESATNNIVSFNDHYASTAWNVLPFSILNIPTWISHLPTGYGDGMGRIMGSAFYDLT
                     EKDFTIIVANLSDPARANTWVQDLNRNAETHTGSTFIIGSDSNDLIQGGQGNDYLEGR
                     AGNDTFRDGGGYNVILGGQGSNSLQLQQSVKDFSFAHDGAGTLYMRDAHGAISITRDI
                     GTLVSKESGLFWGLLKDEVSHSVTANGLVAGANLTAYASSIKGSAANDTLVAGAKGDW
                     LFGQDGNDVLVGGAGNDTFVGGAGNDQMQAGGGSDTFLFNGAFGHDRISGYDAGDKLV
                     FLGVPGAGAGHDYSQYLSQVGNDTLLKVGDSSVTLVGVGLGQVGGEGIAFALI"
BASE COUNT     2430 a   3758 c   3812 g   2152 t
ORIGIN

  Query Match             2.2%;  Score 71.6;  DB 1;  Length 12152;
  Best Local Similarity   61.1%;  Pred. No. 1.6e-05;
  Matches  116;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy       1224 CGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATA 1283
              ||| || ||| |  |||| |||    ||  ||||||  | |||   || || || ||     |
Db       8121 CGGCACCTCCATGGCCGCGCCCCATGTTGCCGGCAGCGTGGCGGTGTTGATGGAGCGTTT 8180

Qy       1284 CCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGG 1343
              |||||  |||| |  || | |  |      | |  |||||  || || |||    ||||| ||
Db       8181 CCCGTACATGACCGGCGCCCAGGTCGCCAGCGTGCTGCGCACCACCGCTACCGACATGGG 8240

Qy       1344 TGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAA 1403
                ||    ||||| ||        || |||||||||  || |   |     ||| |||||||    |
Db       8241 CGCGCCGGGCGTCGATGCGTTGTACGGCTGGGGCATGATCAACCTGGGCAAGGCCATCGA 8300

Qy       1404 CGGACCCGCG 1413
              ||| ||  ||
Db       8301 CGGTCCATCG 8310


RESULT 35
AE003939
LOCUS       AE003939               11186 bp    DNA     linear   BCT 15-JUN-2001
DEFINITION  Xylella fastidiosa 9a5c, section 85 of 229 of the complete genome.
ACCESSION   AE003939 AE003849
VERSION     AE003939.1  GI:9105949
KEYWORDS    .
SOURCE      Xylella fastidiosa 9a5c.
  ORGANISM  Xylella fastidiosa 9a5c
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xylella.
REFERENCE   1  (bases 1 to 11186)
  AUTHORS   Simpson,A.J., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
            Alvarenga,R., Alves,L.M., Araya,J.E., Baia,G.S., Baptista,C.S.,
            Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.,
            Bueno,M.R., Camargo,A.A., Camargo,L.E., Carraro,D.M., Carrer,H.,
            Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C., Costa-Neto,C.M.,
            Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorry,H.,
            Facincani,A.P., Ferreira,A.J., Ferreira,V.C., Ferro,J.A.,
            Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R.,
            Garnier,M., Goldman,G.H., Goldman,M.H., Gomes,S.L., Gruber,A.,
            Ho,P.L., Hoheisel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P.
            and Marino,C.L.
  TITLE     The genome sequence of the plant pathogen Xylella fastidiosa. The
            Xylella fastidiosa Consortium of the Organization for Nucleotide
            Sequencing and Analysis
  JOURNAL   Nature 406 (6792), 151-157 (2000)
  MEDLINE   20365717
   PUBMED   10910347
REFERENCE   2  (bases 1 to 11186)
  AUTHORS   Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
            Alvarenga,R., Alves,L.M.C., Araya,J.E., Baia,G.S., Baptista,C.S.,
            Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.S.,
            Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M.,
            Carrer,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R.,
            Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E.,
            Docena,C., El-Dorry,H., Facincani,A.P., Ferreira,A.J.S.,
            Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C.,
            Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S.,
            Gomes,S.L., Gruber,A., Ho,P.L., Hoheisel,J.D., Junqueira,M.L.,
            Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuramae,E.E., Laigret,F.,
            Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A.,
            Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N.,
            Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L.,
            Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C.,
            Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A.,
            Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G.,
            Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C.,
            Palmieri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira
            Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V.,
            de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V.,
            Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva
            Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de
            Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M.,
            Tsuhako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S.,
            Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and
            Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP
```

```
               13083-970, Brazil
FEATURES                Location/Qualifiers
     source             1. .11186
                        /organism="Xylella fastidiosa 9a5c"
                        /db_xref="taxon:160492"
                        /clone="9a5c"
     gene               complement(128. .631)
                        /gene="XF1012"
     CDS                complement(128. .631)
                        /gene="XF1012"
                        /note="similar to SP|P30134 (percent identity: 54 %/query
                        alignment coverage:  88.6 %/subject alignment coverage:
                        83.1 %); identified by sequence similarity; putative; ORF
                        located using Glimmer/RBSfinder"
                        /codon_start=1
                        /transl_table=11
                        /product="conserved hypothetical protein"
                        /protein_id="AAF83822.1"
                        /db_xref="GI:9105950"
                        /translation="MTATPPPPALPEHIHWMRHALTLAHRAATEFDEIPVGAVLISPE
                        GTLLGEGWNYNITSHDPSAHAEIMALRAAGHQLRNHRMPGCTLYVTLEPCLMCAMAII
                        HARIAHLIYAAADPKTGACGSTFDILNDPRHNHHVHVYGGLLAEEASRRLTNYFRTKR
                        GQPPLPQ"
     gene               complement(639. .1121)
                        /gene="XF1013"
     CDS                complement(639. .1121)
                        /gene="XF1013"
                        /note="similar to GI|1787038 (percent identity: 61 %/query
                        alignment coverage:  90.0 %/subject alignment coverage:
                        92.9 %); identified by sequence similarity; putative; ORF
                        located using Glimmer/RBSfinder/Start codon shift: 12"
                        /codon_start=1
                        /transl_table=11
                        /product="conserved hypothetical protein"
                        /protein_id="AAF83823.1"
                        /db_xref="GI:9105951"
                        /translation="MNTRHTIPKHENTQTQTAPLIDAKIHVASFRQVRQARRDELIKD
                        YVELISDLLADSGQARQVDIAAHLGVAQPTVAKMLKRLARDGWIIQRPYRGTFLTPAG
                        KALAAASRECHQTVERFLLALGIDAATARRDAEGIEHHVSETTLAAFRAFLHRHSDTT
                        "
     gene               1102. .1284
                        /gene="XF1014"
     CDS                1102. .1284
                        /gene="XF1014"
                        /note="hypothetical protein; identified by sequence
                        similarity; putative; ORF located using
                        Glimmer/RBSfinder/Start codon shift: -12"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein"
                        /protein_id="AAF83824.1"
                        /db_xref="GI:9105952"
                        /translation="MVCLVFIGDCIVGHRYFGVTFLCLRPSHGESALGVPWDVLNGEM
                        LVSGVWLCLVWAAIFL"
     gene               1286. .2704
                        /gene="XF1015"
     CDS                1286. .2704
                        /gene="XF1015"
                        /note="similar to GI|5670182 (percent identity: 64 %/query
                        alignment coverage:  93.6 %/subject alignment coverage:
                        100.9 %); identified by sequence similarity; putative; ORF
                        located using Glimmer/RBSfinder"
                        /codon_start=1
                        /transl_table=11
                        /product="manganese transport protein"
                        /protein_id="AAF83825.1"
                        /db_xref="GI:9105953"
                        /translation="MSCGIANAIRWDVSLLIHRSMSSSMSSVQPTSPVSDSPSLGEMH
                        ASVAVSRRGHWGFRLLAFLGPGYMVSVGYMDPGNWATGLAGGSRFGYMLLSVILLSNV
                        MAIVLQALAARLGIASDMDLAQACRARYSRGTTLALWVVCELAIIACDLAEVIGTAIA
                        LNLLLGVPIIWGVVITAVDVVLVLLLMHRGFRALEAFVIALLLVIFGCFVVQIVLAAP
                        PLQEVLGGFVPRWQVVADPQALYLAIGIVGATVMPHNLYLHSSIVQTRAYPRTPVGRR
                        SALRWAVADSTLALMLALFINASILILAAAVFHAQHHFDVEEIEQAYQLLAPVLGVGV
                        AATLFATALLASGINSTVTATLAGQIVMEGFLRLRLRPWLRRVLTRGLAIVPVIVVVA
                        LYGEQGTGRLLLLSQVILSMQLPFAVIPLLRCVADRKVMGALVAPRWLMVVAWLIAGV
                        IVVLNVKLLGDYAVHLMVGVSD"
     gene               2720. .2884
                        /gene="XF1016"
     CDS                2720. .2884
                        /gene="XF1016"
                        /note="hypothetical protein; identified by sequence
                        similarity; putative; ORF located using Glimmer/RBSfinder"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein"
                        /protein_id="AAF83826.1"
                        /db_xref="GI:9105954"
                        /translation="MIRQGRCSQWSWCFVLMHRFAVFQGAVYGMPVCMVVDGGSRRHN
                        DVLAICWVSF"
     gene               complement(2862. .3035)
                        /gene="XF1017"
     CDS                complement(2862. .3035)
                        /gene="XF1017"
                        /note="hypothetical protein; identified by sequence
                        similarity; putative; ORF located using Glimmer/RBSfinder"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein"
                        /protein_id="AAF83827.1"
                        /db_xref="GI:9105955"
                        /translation="MNRTRKKDSTTNELFQPQPQRNLTTSQNITELSAPAPQPKQIAS
                        WHDASHIKTIPNK"
     gene               complement(3202. .3966)
                        /gene="XF1018"
     CDS                complement(3202. .3966)
                        /gene="XF1018"
                        /note="similar to GI|6321421 (percent identity: 25 %/query
                        alignment coverage:  95.7 %/subject alignment coverage:
                        48.3 %); identified by sequence similarity; putative; ORF
                        located using Glimmer/RBSfinder/Start codon shift: -99"
                        /codon_start=1
                        /transl_table=11
                        /product="arginine-tRNA-protein transferase"
                        /protein_id="AAF83828.1"
                        /db_xref="GI:9105956"
                        /translation="MAIDSKPHDDLQLFKTNLHPCGYWPDRWASDLVMDPNDPRLGAI
                        YPQALAWGFRRSGNLLYRPHCEHCNACVPVRVNVNAFVPNRSQRRCLARNATLVTRIV
                        PAERNAEQLSLYRRYLHQRHPDGGMDGHGAIEFDQFLIGPWGYGRFMEIREPATNGTP
                        GQLLAVAVTDLTHQALSAVYTFYEPNAAARGLGTLAILHQIHWAQREQRPYLYLGYWI
                        KDHFKMDYKRRFQKLEIYDGYRWRPFSTTYPTTHTL"
     gene               4148. .4267
                        /gene="XF1019"
     CDS                4148. .4267
                        /gene="XF1019"
                        /note="hypothetical protein; identified by sequence
                        similarity; putative; ORF located using Glimmer/RBSfinder"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein"
                        /protein_id="AAF83829.1"
                        /db_xref="GI:9105957"
                        /translation="MLRAITKGLFALLGVNAGFVVVLGQRVGCEGVLDVWMML"
     gene               complement(4388. .4879)
                        /gene="XF1020"
     CDS                complement(4388. .4879)
                        /gene="XF1020"
                        /note="similar to GI|155377 (percent identity: 48 %/query
                        alignment coverage: 101.8 %/subject alignment coverage:
                        98.8 %); identified by sequence similarity; putative; ORF
                        located using Glimmer/RBSfinder"
                        /codon_start=1
                        /transl_table=11
                        /product="pathogenicity-related protein"

  Query Match             2.2%;  Score 71.2;  DB 1;  Length 11186;
  Best Local Similarity   51.2%;  Pred. No. 2e-05;
```

```
  Matches  250;  Conservative    0;  Mismatches  223;  Indels   15;  Gaps    3;

Qy    1215 TCAAATTGCCGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCT 1274
            | || |    || || || ||| | || ||  | ||  | ||  | |||||  || |
Db    9150 TGAACTACAGGGCACCTCGTTTGCAGCGCCGTTGGTTTCTGGTGCAGCGGCATTGGTATG 9209

Qy    1275 GCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCA 1334
            | ||  ||  || |    | |||||||||     ||||      || |||||     |||||
Db    9210 GGAGGCATTTCCCTATTTCAGCAACGACCTGGTGCGGCAGACCCTGCTTGGTACGGCGAC 9269

Qy    1335 GGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAA 1394
             ||  | || |      ||||||||          ||||||| ||      ||| ||| ||||||
Db    9270 TGATCTTGGCGTCCCCGGCGTGGATGAAGTCTTCGGCTATGGTGCACTGAATGTGGGTAA 9329

Qy    1395 GGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAGGTAC 1454
             || ||    || || ||      ||      |  || ||  |      | |           |
Db    9330 AGCAGTGCTGGGCCCGGCTAAGTTCGATTGGGGTGATGTCCAAGTC---------AGCTT 9380

Qy    1455 ATCCGATATTGCCTACTCCTTCCGTAACGACATTTCAGGCACGGGCGGCCTGATCAAAAA 1514
                 || |      | | |||   | || |||||  | ||||   |||||| |||||||
Db    9381 CGATGACAGGCGTTCCACCTGGGGCAATGACATCACCGGCAGCGGCGGCTTGATCAAGCG 9440

Qy    1515 AGGCGGCAGCCAACTGCAACTGCACGGCAACAACA---CCTATACGGGCAAAACCATTAT 1571
             ||    | |    ||||    ||      |||| ||| |          || || ||      |||      |
Db    9441 TGGTACCGGAACACTGGTGCTTGGTGGCAGCAATAATCAATACACCGGAGCCACCCAAGT 9500

Qy    1572 CGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAA 1631
            |   || ||| |   ||   |   || | ||             ||| |     |   | || |   | |
Db    9501 CCTGGGAGGTACATTGCAGGCGTCCAGCTTGGGTGTATCCGCAGTCAGTGTTGCTAACGC 9560

Qy    1632 AGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGTAGCCTGAACAGCGACGGCATTGT 1691
             |    || |||||  || |||               | |||||||      |||| ||   | |||||    |
Db    9561 TGCGACGTTGATTGGTAGCGG---TCGCTTCGGCGGTGCGCTGAGCAATGCCGGCACCCT 9617

Qy    1692 CTATCTGG 1699
              | || |
Db    9618 GGAGCTTG 9625


RESULT 36
AL646057
LOCUS       AL646057               215050 bp    DNA     linear   BCT 07-DEC-2001
DEFINITION  Ralstonia solanacearum GMI1000 chromosome, complete sequence;
            segment 1/19.
ACCESSION   AL646057 AL646052
VERSION     AL646057.1  GI:17427008
KEYWORDS    .
SOURCE      Ralstonia solanacearum.
  ORGANISM  Ralstonia solanacearum
            Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
            Ralstonia.
REFERENCE   1  (bases 1 to 215050)
  AUTHORS   Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,
            Arlat,M., Billault,A., Brottier,P., Camus,J.C., Cattolico,L.,
            Chandler,M., Choisne,N., Claudel-Renard,C., Cunnac,S., Demange,N.,
            Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schiex,T.,
            Siguier,P., Thebault,P., Whalen,M., Wincker,P., Levy,M.,
            Weissenbach,J. and Boucher,C.A.
  TITLE     Genome sequence of the plant pathogen Ralstonia solanacearum
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 215050)
  AUTHORS   Boucher,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
            Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
            Moleculaire des Interactions Plantes-Microorganismes INRA-CNRS,
            BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
            Dausset-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, LMGM CNRS
            118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
            URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
            Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
            F31326 Castanet-Tolosan Cedex. Laboratoire de Genetique Cellulaire
            INRA, BP27, F31326 Castanet-Tolosan Cedex
COMMENT     Christian.Boucher@toulouse.inra.fr
            http://sequence.toulouse.inra.fr/R.solanacearum.html.
FEATURES             Location/Qualifiers
     source          1..215050
                     /organism="Ralstonia solanacearum"
                     /strain="GMI1000"
                     /db_xref="taxon:305"
     rep_origin      1..11
                     /note="oriC or RS01885
                     predicted by FrameD"
                     /evidence=not_experimental
     gene            486..620
                     /gene="rpmH"
                     /note="RSc0001; RS06178"
     CDS             486..620
                     /gene="rpmH"
                     /function="structural elements; ribosome constituents;
                     ribosomal proteins - synthesis, modification"
                     /note="Product confidence : probable
                     Gene name confidence : probable
                     predicted by Homology"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="PROBABLE 50S RIBOSOMAL PROTEIN L34"
                     /protein_id="CAD13529.1"
                     /db_xref="GI:17427009"
                     /translation="MKRTYQPSVTRRKRTHGFRVRMKTRGGRAVLNARRAKGRKRLAI
                     "
     gene            718..1113
                     /gene="rnpA"
                     /note="RSc0002; RS01824"
     CDS             718..1113
                     /gene="rnpA"
                     /EC_number="3.1.26.5"
                     /function="macromolecule metabolism; macromolecule
                     degradation; degradation of rna"
                     /note="Product confidence : probable
                     Gene name confidence : probable
                     predicted by Homology
                     predicted by FrameD"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="PROBABLE RIBONUCLEASE P PROTEIN COMPONENT"
                     /protein_id="CAD13530.1"
                     /db_xref="GI:17427010"
                     /translation="MKTDEFSSVFALRPVRRSRHFVLYVRANGHPQARLGIVIGKKFA
                     RRAVERNLIKRQCRELFRLRQAALGGRDVLIRLQTKFPREDVPTVAAFKRLCREELSY
                     LFEIAARPLPAPPVPCPAAAPVPSDGAAP"
     gene            1110..1424
                     /gene="RSc0003"
                     /note="RS01825"
     CDS             1110..1424
                     /gene="RSc0003"
                     /function="miscellaneous; hypothetical/global homology"
                     /note="Product confidence : hypothetical
                     Gene name confidence : hypothetical
                     predicted by Codon_usage
                     predicted by Homology
                     predicted by FrameD"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="CONSERVED HYPOTHETICAL PROTEIN"
                     /protein_id="CAD13531.1"
                     /db_xref="GI:17427011"
                     /translation="MTRVLLFLLRVYKVAFSPFVGAQCRFLPTCSDYARDAVLTHGPA
                     IGSYLAAKRLCRCHPFAQGGYDPVPPAAGDAALRSTDSASTETTAADAPAARPSIHLP
                     RP"
     gene            1431..3092
                     /gene="RSc0004"
                     /note="RS01826"
```

```
CDS             1431. .3092
                /gene="RSc0004"
                /function="miscellaneous; hypothetical/global homology"
                /note="Product confidence : probable
                Gene name confidence : hypothetical
                predicted by Codon_usage
                predicted by Homology
                predicted by FrameD"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="PROBABLE TRANSMEMBRANE PROTEIN"
                /protein_id="CAD13532.1"
                /db_xref="GI:17427012"
                /translation="MDIKRTILWVIFSLSVVLLFDNWQRANGHQSMFFPTPQTVTTTA
                AAPGGTPAGDVPKAAAPAAAGSQAAPATGAVSQTPASEKIVVTTDVIRATVDTAGAIV
                TKLELLTQKDHDGNPMVLFDRSLERTYLARSGLIGGDFPNHTTVFTASAGPRDLGTGG
                EVSLTLTADKGGAKLAKTYVFKRGSYVIDTRFDVTNDGAAPINPTLYMELARDGGAVE
                QSRFYSTFTGPAVYTDTDHYHKITFADIDKSKAHVPAPTDSGWVAMVQHYFASAWIPA
                ASAKREFYVDRIDTNFYRVGMQQALGTVAPGASVSATARLFAGPQEERMLEGITPGLE
                LVKDYGWLTIIAKPLFWLLEKIHKLLGNWGWSIVALTVLVKLVFFPLSATSYRSMAKM
                KDLQPRMTAIRERHKGDPQKMNQEMMTLYRTEKVNPLGGCLPIVIQIPVFIALYWVLL
                SSVEMRGAPWLGWVHDLASPDPFYILPILMAVSMFVQTRLNPTPPDPVQAKMMMFMPI
                AFSVMFFFFPAGLVLYWVVNNCLSIAQQWSINRMLGTNKKAAAAK"
gene            3222. .4667
                /gene="thdF"
                /note="RSc0005; RS01827"
CDS             3222. .4667
                /gene="thdF"
                /function="cell processes; protection responses;
                detoxification"
                /note="Product confidence : probable
                Gene name confidence : probable
                predicted by Codon_usage
                predicted by Homology
                predicted by FrameD"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="PROBABLE THIOPHENE AND FURAN OXIDATION PROTEIN"
                /protein_id="CAD13533.1"
                /db_xref="GI:17427013"
                /translation="MTASSHAMTSPTVSDAASAPIRTVPIAAIATAPGRGGIGVVRVS
                GPDVRAVMQAVCGRLLPPRQATYLPFLDADGAAIDRGIALWFPAPHSYTGEDVLELQG
                HGGPVVMQLLLSRCLRAGHGIGLRVAEPGEFTRRAFLNDKLDLAQAEAVADLIEASTE
                AAARSAARSLDGVFSQTVHALVERVIHLRMLVEATLDFPEEEIDFLEAADARGQLADI
                RARLDGVLAQARQGALLREGLHVVLAGQPNVGKSSLLNALAGAELAIVTPIAGTTRDK
                VQQTIQIEGIPLNIVDTAGLRDTEDEVERIGIERTWAAIARADVVLHLLDAADYRAHG
                LSAEDAAIDARIAEHVPPGVPTLRVINKIDLAGAAVPGRVDAQPPEVWLSARDGSGIE
                LLRAALLEIAGWQGGGEGLYLARERHLSALRSAREHLTIAADHADQRAQSLDLFAEEL
                RLAQEALNSITGAFSSDDLLGVIFSRFCIGK"
gene            4757. .5080
                /gene="RSc0006"
                /note="RS01828"
CDS             4757. .5080
                /gene="RSc0006"
                /function="miscellaneous; hypothetical/global homology"
                /note="Product confidence : hypothetical
                Gene name confidence : hypothetical
                predicted by Codon_usage
                predicted by FrameD"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="CONSERVED HYPOTHETICAL PROTEIN"
                /protein_id="CAD13534.1"
                /db_xref="GI:17427014"
                /translation="MRALFICSRNRLRSPTAESIFAHWPNVDTDSAGLAPDADVPLTA
                DQLEWAEVIFVMERSHRRRLSQRFGPWLRGKRVVCLDIPDDYLFMQPELVALLERRAG
                PYLRS"
gene            5188. .5721
                /gene="RSc0007"
                /note="RS01829"
CDS             5188. .5721
                /gene="RSc0007"
                /function="miscellaneous; unknown"
                /note="Product confidence : hypothetical
                Gene name confidence : hypothetical
                predicted by FrameD"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="HYPOTHETICAL PROTEIN"
                /protein_id="CAD13535.1"
                /db_xref="GI:17427015"
                /translation="MGRGNDYTWWRKAGEAAEALFYANGWAARAAYALGLQGRLRVDR
                RDVYLPLKPLAELLRVAFASDFHAGPLTDPRLLDAVVRRIDAFAPHILLLGGDFVSLH
                HRHVSSLAVRLGELRVPAGMFGVYGKRGKPEWAVWRGRAFKAREHCRYADRHAARLPD
                YAVPPAPRRCIRRRTAR"
gene            complement(6204. .6416)
                /gene="rpsU"
                /note="RSc0008; RS01830"
CDS             complement(6204. .6416)
                /gene="rpsU"
                /function="structural elements; ribosome constituents;

  Query Match             2.2%;  Score 70.8;  DB 1;  Length 215050;
  Best Local Similarity   51.3%;  Pred. No. 3e-05;
  Matches  193;  Conservative    0;  Mismatches  177;  Indels    6;  Gaps    1;

Qy       1255 GGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACC 1314
              ||| | | ||  ||||||      |    | |||||    |||||  |||| | |   ||
Db     154091 GGCGCCGAGGTGCTGCTGGAAACGCGCTTCCCGTACCTGAGCGCCGACCAGCGCCGCGTG 154150

Qy       1315 ACGCTGCTGACAACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGG 1374
                |||| ||| || |  | || |  |||         |   |||||  |       ||||||
Db     154151 GTGCTGAAGACCACCGAGCTGGCCTCCGGCTACCCGGTGATGGACGACGCCGAAGGCTGG 154210

Qy       1375 GGACTGCTGGATGCGGGTAAGGCCATGAACGGACCCGCGTCCTTTCCGTTCGGCGACTTT 1434
              || |  ||| |    |      ||||    ||||     |    | ||        | |||    |
Db     154211 GGCCGCCTGAACCTGTTCGCGGCCGCCGACGGCTATGGCGCGTTCAACGGCAGCGTGGTC 154270

Qy       1435 ------ACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCCTTCCGTAACGACATT 1488
                    |   |     || | ||        | | ||    ||| || |||  || ||||   ||
Db     154271 GTGTCGATGGATGCCTCGCAGGGCGGCTTCAATGCCGCCGACACCTGGCGCAACGCGGTT 154330

Qy       1489 TCAGGCACGGGCGGCCTGATCAAAAAAGGCGGCAGCCAACTGCAACTGCACGGCAACAAC 1548
               | ||| |||||    ||||        | |||  | ||     ||||   |||   ||||| ||||
Db     154331 GCCGGCGCGGGCAAGCTGACGCTGCAGGGCACCGGCACGCTGCGCCTGGCCGGCAGCAAC 154390

Qy       1549 ACCTATACGGGCAAAACCATTATCGAAGGCGGTTCGCTGGTGTTGTACGGCAACAACAAA 1608
              ||||| ||||||    ||      |||   |||||   ||||| |      ||       || |
Db     154391 ACCTACACGGGCGGCACGCAGGTCGCGGGCGGCGTGCTGGGGGCCGACTCGGCCACCGCG 154450

Qy       1609 TCGGATATGCGCGTCG 1624
               |  |  | | ||| ||
Db     154451 TTCGGCACGGGCGACG 154466


RESULT 37
SMASSPH
LOCUS       SMASSPH                 7030 bp    DNA     linear   BCT 03-OCT-1998
DEFINITION  Serratia marcescens DNA for SSP-h1 and SSP-h2, complete cds.
ACCESSION   D78380
VERSION     D78380.1  GI:3688584
KEYWORDS    SSP-h2; SSP-h1.
SOURCE      Serratia marcescens (strain:IFO3046) DNA.
  ORGANISM  Serratia marcescens
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Serratia.
REFERENCE   1  (sites)
  AUTHORS   Ohnishi,Y., Beppu,T. and Horinouchi,S.
  TITLE     Two genes encoding serine protease homologues in Serratia
            marcescens and characterization of their products in Escherichia
            coli
  JOURNAL   J. Biochem. 121 (5), 902-913 (1997)
```

```
  MEDLINE   97335937
REFERENCE   2  (bases 1 to 7030)
  AUTHORS   Ohnishi,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-1995) Yasuo Ohnishi, The University of Tokyo,
            Department of Biotechnology; Yayoi 1-1-1, Bunkyo-ku, Tokyo
            113-8467, Japan (E-mail:ayasuo@hongo.ecc.u-tokyo.ac.jp,
            Tel:+81-3-3812-2111, Fax:+81-3-3812-0544)
COMMENT     On Oct 3, 1998 this sequence version replaced gi:1100762.
            Sequence updated (24-Sep-1998).
FEATURES             Location/Qualifiers
     source          1..7030
                     /organism="Serratia marcescens"
                     /strain="IFO3046"
                     /db_xref="taxon:615"
     CDS             572..3682
                     /codon_start=1
                     /transl_table=11
                     /product="SSP-h1"
                     /protein_id="BAA33455.1"
                     /db_xref="GI:3688585"
                     /translation="MEKKLRGVKSARAAFGWPEIGVHRAQWSAMLICLTAIGGAQASS
                     YIENGKAGDPASWRSSEFNAEWGLGAIHADQAYAAGYTGKGIKLGIFDQPVYAKHPEF
                     AGENKVINLVTEGIREYTDPYIPVKKGDAFRYDGTPSVDSDGTLGSHGTHVGGIAAGS
                     RDGGAMHGVAFNAQIISAENGDPGPEDGIILGNDGAVYQAGWDALVASGARIINNSWG
                     IGITEKFDEGGYDPAYPHFTVNDAQKQFDQIKQILGTKPGGAYQGAIDAARSGVVTIF
                     AAGNDGNLNNPDAMAGLAYFVPDIAPNWLSVASLQDPTNTGDYSISTFSSRCGYTASL
                     CVSAPGTRVYSSIIEGTSVENLTTGYAKYSGTSMAAPHVAGSVAVLMERFPYLNGAQV
                     AEVLKTTATDMGAPGIDALYGWGMINLGKAINGPGMLVTAEDIPEEFRIPDPTGVAYG
                     PTQFVVDLPGVGAVLDKGKPTERVCSDVLCGLDFWSNDISGHGGLTKQGIGTLVLTGN
                     NTYSGPTLVNQGRLAINGSVTSAVSVQSGGIVGGSGTVGSLTARRGGTVAPGNSIGTL
                     NVAGDVSFEPGSRYAVEVGPNGQSDRIQSGGAATIGGGEVAVTLENSSNLLTQSEVRS
                     LLGQQYTILSAQQGVSGQFDAVAPNYLFLGTGLSYQPTGVTLSVGRNGTSFASVAQTP
                     NERAVAAAADALAAGNPVYESLLNSGTAGEARQAFRQLSGQIHADIASALVNDSRYLR
                     EALNGRLRQAEGLASSSAIKADEDGAWAQLLGAWDHASGDANATGYQASTYGVLVGLD
                     SAAAADWRLGVATGYTRTSLHGGYGSKADSDNYHLAAYGDKQFGALALRGGAGYTWHR
                     IDTKRSVNYGMQSDRDTAKYSARTEQLFAEAGYSVKGEWLNLEPFVNLAYVNFENNGI
                     AESGGAAALRGDKQHTDATVSTLGLRADTEWQVSPGTTVALRSELGWQHQYGGLERGT
                     GLRFNGGNAPFVVDSVPVSRDGMVLKAGAEVAVNENASLSLGYGGLLSQNHQDNSVNA
                     GFTWRF"
     CDS             3801..6905
                     /codon_start=1
                     /transl_table=11
                     /product="SSP-h2"
                     /protein_id="BAA11383.1"
                     /db_xref="GI:1100764"
                     /translation="MTTTMGFDSKGNARRSPWRLNALVCCMALAGYAQAAPHEVSGQA
                     GDPASWRSAEFNANWGLGAIHADEAYAAGYTGKGQKVGIFDTPVNRHPEFAGDGKLIN
                     VVTEGYRAYTDPHRPGINAGDRFYFDGTFHFYSGSQGMLSNHGVHVAGISAANRDGVG
                     MHGVAFDSQVISVDNDNDGPAYGEFLGLDGAVTNAGWQAMIKSGVRVINNSWGVSIPD
                     FLSDGGRDPNALHFELKDAQEQFDQVKPLLGSLAGAGYQGAIDAARKNILVLFAAGND
                     GNYNQPDVISGLAYFVPDIAPNWLSVASVAQDATSTNSVPYTISSFSSRCGYTASFCV
                     SSPGSKIYSTVANGSDPANLVSDYGNKNGTSMATPHVTGAVAVLLQRFPYMSSAQIAD
                     VLKTTATDMGAPGIDALYGWGMINLGKAINGPGMFYTVEDIPAEFRIPDPTGVAYGPT
                     QFVANIPGLGAEVDAGTLHARKCDDISCELEVYSNNISGHGGLTKEGRGTLVMTGNNT
                     YSGPTLVNQGRLAINGSVTSAVSVQSGGIVGGSGTVGSLIARRGGTVAPGNSIGTLNV
                     AGDVSFEPGSRYAVEVGPNGQSDRIQSGGAATIGGGEVAVTLENSSNLLTQSEVRSLL
                     GQQYTILSAQQGVSGQFDAVAPNYLFLGTGLSYQPTGVTLSVGRNGTSFASVAQTPNE
                     RAVAAAADALAAGNPVYESLLNSGTAGEARQAFRQLSGQIHADIASALVNDSRYLREA
                     LNGRLRQAEGLASSSAIKADEDGAWAQLLGAWDHASGDANATGYQASTYGVLVGLDSA
                     AAADWRLGVATGYTRTSLHGGYGSKADSDNYHLAAYGDKQFGALALRGGAGYTWHRID
                     TKRSVNYGMQSDRDTAKYSARTEQLFAEAGYSVKGEWLNLEPFVNLAYVNFENNGIAE
                     SGGAAALRGDKQHTDATVSTLGLRADTEWQVSPGTTVALRSELGWQHQYGGLERGTGL
                     RFNGGNAPFVVDSVPVSRDGMVLKAGAEVAVNENASLSLGYGGLLSQNHQDNSVNAGF
                     TWRF"
BASE COUNT     1370 a   2099 c   2431 g   1130 t
ORIGIN

  Query Match             2.1%;  Score 68;  DB 1;  Length 7030;
  Best Local Similarity   57.5%;  Pred. No. 0.0001;
  Matches  122;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy       1224 CGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATA 1283
              ||| || ||  |  |  | ||   ||| |||||| ||| |||  || ||||||||   |
Db       4886 CGGCACCTCGATGGCGACGCCGCACGTCACCGGCGCGGTGGCGGTGTTGCTGCAGCGCTT 4945

Qy       1284 CCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGG 1343
              |||||  ||||||   |   |  |      |  ||||   || || ||    ||||| ||
Db       4946 CCCGTACATGAGCTCGGCGCAGATTGCCGACGTGCTGAAAACCACCGCCACCGACATGGG 5005

Qy       1344 TGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAA 1403
               ||    ||| | |||      | |||||||||| || |  |   ||| |||||||| ||
Db       5006 CGCGCCGGGCATCGACGCGCTCTACGGCTGGGGAATGATCAACCTGGGCAAGGCCATCAA 5065

Qy       1404 CGGACCCGCGTCCTTTCCGTTCGGCGACTTTA 1435
              ||| || |     ||      || |||   ||
Db       5066 CGGCCCGGGCATGTTCTATACCGTCGAAGATA 5097


RESULT 38
AE004006
LOCUS       AE004006            11476 bp    DNA     linear   BCT 15-JUN-2001
DEFINITION  Xylella fastidiosa 9a5c, section 152 of 229 of the complete genome.
ACCESSION   AE004006 AE003849
VERSION     AE004006.1  GI:9106932
KEYWORDS    .
SOURCE      Xylella fastidiosa 9a5c.
  ORGANISM  Xylella fastidiosa 9a5c
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xylella.
REFERENCE   1  (bases 1 to 11476)
  AUTHORS   Simpson,A.J., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
            Alvarenga,R., Alves,L.M., Araya,J.E., Baia,G.S., Baptista,C.S.,
            Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.,
            Bueno,M.R., Camargo,A.A., Camargo,L.E., Carraro,D.M., Carrer,H.,
            Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C., Costa-Neto,C.M.,
            Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorry,H.,
            Facincani,A.P., Ferreira,A.J., Ferreira,V.C., Ferro,J.A.,
            Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R.,
            Garnier,M., Goldman,G.H., Goldman,M.H., Gomes,S.L., Gruber,A.,
            Ho,P.L., Hoheisel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P.
            and Marino,C.L.
  TITLE     The genome sequence of the plant pathogen Xylella fastidiosa. The
            Xylella fastidiosa Consortium of the Organization for Nucleotide
            Sequencing and Analysis
  JOURNAL   Nature 406 (6792), 151-157 (2000)
  MEDLINE   20365717
   PUBMED   10910347
REFERENCE   2  (bases 1 to 11476)
  AUTHORS   Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
            Alvarenga,R., Alves,L.M.C., Araya,J.E., Baia,G.S., Baptista,C.S.,
            Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.S.,
            Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M.,
            Carrer,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R.,
            Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E.,
            Docena,C., El-Dorry,H., Facincani,A.P., Ferreira,A.J.S.,
            Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C.,
            Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S.,
            Gomes,S.L., Gruber,A., Ho,P.L., Hoheisel,J.D., Junqueira,M.L.,
            Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuramae,E.E., Laigret,F.,
            Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A.,
            Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N.,
            Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L.,
            Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C.,
            Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A.,
            Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G.,
            Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C.,
            Palmieri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira
            Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V.,
            de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V.,
            Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva
            Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de
            Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M.,
            Tsuhako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S.,
            Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.
  TITLE     Direct Submission
```

```
JOURNAL   Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and
          Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP
          13083-970, Brazil
FEATURES             Location/Qualifiers
     source          1. .11476
                     /organism="Xylella fastidiosa 9a5c"
                     /db_xref="taxon:160492"
                     /clone="9a5c"
     gene            69. .167
                     /gene="XF1847"
     CDS             69. .167
                     /gene="XF1847"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF84653.1"
                     /db_xref="GI:9106933"
                     /translation="MAVVLHELGRIERTLFILNWLQNVELRRSTSI"
     gene            complement(309. .1736)
                     /gene="XF1848"
     CDS             complement(309. .1736)
                     /gene="XF1848"
                     /note="similar to SP|P06713 (percent identity: 51 %/query
                     alignment coverage:  96.8 %/subject alignment coverage:
                     98.1 %); identified by sequence similarity; putative; ORF
                     located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="two-component system, regulatory protein"
                     /protein_id="AAF84654.1"
                     /db_xref="GI:9106934"
                     /translation="MVEIDATSPRIWVVDDDRSVRFVLSTALRDAGYIVEGFESTCAA
                     LQALAASPIPDLLFTDVRMPGDDGLILLDKLKARHPHLPVIVMSAYTDVASTAGAFRS
                     GAHEFLSKPFDLDDAVALAARVLPHTEGSHGGPPILPAANSNPVLIGSTPVMRALFRA
                     IGRLAQAPLSVLINGETGTGKELVARALHNESPRAQAPFVALNAAAIPTELLESELFG
                     HEAGAFTGATKRHIGRFEQADGGTLFLDEIGDMPLSLQTRLLRVLAENEFFRVGGREL
                     IRVNVRVITATHQDLDMLVEQGRFRSDLLHRLDVVRLQLPSLRERRDDVAQLAENFLA
                     MAGHKLDISVKRLSADALEILRRYDWPGNVRELENICWRLAALVNTETIDANDVEAVL
                     THGKRKRCNTESDWDIHLAKWAEGRLAAGATAVHAEARKCLDRTLLDVALRLTQGRRA
                     EAAARLGLGRNTLTRKLGSSQRRRG"
     gene            complement(1729. .2784)
                     /gene="XF1849"
     CDS             complement(1729. .2784)
                     /gene="XF1849"
                     /note="similar to GI|310876 (percent identity: 37 %/query
                     alignment coverage:  78.3 %/subject alignment coverage:
                     76.2 %); identified by sequence similarity; putative; ORF
                     located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="two-component system, sensor protein"
                     /protein_id="AAF84655.1"
                     /db_xref="GI:9106935"
                     /translation="MRMPVPPPSLENLCTPVAWGNAEGRVQGVNAAFAYWLGVSAKRL
                     QQQTLEALEQQGNALSSFLNNDEHNVLCLHRFALGIPGGTPSFAEGWLTRLEAGAWLL
                     EAHPIKECSALDTLQALPNALSAAFKGLAHELRNPLSGLKGAAQLLARRVKHRDQEEC
                     DLVELIGAEIERLNTLLERLLFSAPARPHAQLNIHTVLERVLPLAETEAGCSVILQRD
                     YDPSIPDILGDADRLTQAVWNLVRNAIQAGASRIILRTRVEHGQRIRDRVNVMSLRLE
                     VIDDGDGVPEALTEHVFLPLVSSRAEGSGLGLPLAQQVAHEHHGMLTFRSQPGQTIFI
                     LSLPQMMIESDKDLHYG"
     gene            complement(3161. .3649)
                     /gene="XF1850"
     CDS             complement(3161. .3649)
                     /gene="XF1850"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF84656.1"
                     /db_xref="GI:9106936"
                     /translation="MPFHVQQDGRSRGEKKLGYAGLCKRLTAWRTDAPVHPLQQSLKD
                     WGRAYANLFTKRADFPRFKKIRYRNSRNVLGTVKNVTVNQSCGKWYVSIQTADEITEP
                     VHPSKLNVEIDADVTQLATQSDGTIYLPVNSFKSRQKSQISVRKDYLHKVTTTISKNP
                     TP"
     gene            3851. .6853
                     /gene="XF1851"
     CDS             3851. .6853
                     /gene="XF1851"
                     /note="similar to SP|P09489 (percent identity: 25 %/query
                     alignment coverage:  99.0 %/subject alignment coverage:
                     94.7 %); identified by sequence similarity; putative; ORF
                     located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="serine protease"
                     /protein_id="AAF84657.1"
                     /db_xref="GI:9106937"
                     /translation="MQTIKNKFVVRTILATTVTIVLSACGGGLVRIDIPRPTPTPTPT
                     PTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTLPSLPPSLAAKI
                     GRHLELTNTEAAHALGLTGSGFIIGMVDTGVHRNNIALQGKVIRSGVYNFYTDYHAPS
                     APDDVDNHGTLVAQLIAGQPVEHWPGGIAPGASIFSARVYPSEEEFEAYIEEQMTSSH
                     IVDTFEAEASKIRLINADMIAAGVRIQNNSWGHSSYAMWTDPKITNLFGDTYRDFVIG
                     SNGLVVFSTGNEKLSQPNGLARLPSLLRAEGSPSAADDLERGWLAVAGLDSFDHPDQL
                     GSYSNHCGVAMHYCLAAPGNVLVIDSNVTTTPIGNDGYEEQRGTSFSAPLVSGAAALV
                     WEVFPYFNNDLVRQTLLGTATDLGAPGVDEVFGYGALNVGKAVRGPAKFDWGDVKVNF
                     NDSRSTWFNDISGGGGLIKDGTGTLVLAGSANQYSGATQVLGGTLQVSSLGESAVSIA
                     HSATLIGSGRFGGAVSNAGTLEIGTDGLKVQGDYTQSETGRLALNVGDQLSVAGHAAL
                     QGGELQVLGKRDYVSLNTSYSVLHADGGLSGTFSGLTWGPAVLLGEGQLHYGANDANV
                     TLQRLDVNAAARALGIVDTVAQASARRVEQAFREIDIQQAQGRGQLSTSFMAAAAALQ
                     QSGSAQMASASLHSLSGQGHVAATASSFDVVDAGRRSLATRLGERPLGERTVTWRQSL
                     GGSGQDGTAGGTFALSGWLLGRDTQLASGDIVGFAFGQSEASPSGSVSGLRGLDRQVQ
                     GRMYVQRSQGPLYVLGQLGFGSFQRRLDRQLQLGHWSDWASSRYRGQFFSGSVETGYH
                     WRRGAWALTPYVGVDQTQLKTDGFREQGGAGFGLQVQAAQAIRSQLLAGIRTELGWRG
                     VKLRGYGEWQQTLRQSGLNPQASFTATSSWTPLVASPWSSRSGGLLGLSMVWPVGMTV
                     GQWNMGVEHFFGARKRSDSLGLRYQLSF"
     gene            complement(7118. .7366)
                     /gene="XF1852"
     CDS             complement(7118. .7366)
                     /gene="XF1852"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF84658.1"
                     /db_xref="GI:9106938"
                     /translation="MIHVLNKGKELKNFAPMGGASENRKQKTKHAKPCSLLIASSLLG
                     LLNPLNSVMPSTGKKMTGAKTLKKTKRIYGFSLLLQCD"
     gene            7770. .7958
                     /gene="XF1853"
     CDS             7770. .7958
                     /gene="XF1853"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF84659.1"
                     /db_xref="GI:9106939"
                     /translation="MHNAAAAPLRAQGNVSFVFDLSGKPLRSFQMGILRQIKLRWLSG
                     HRYLCFEISFLKIVVIHG"
     gene            7967. .9565
                     /gene="XF1854"
     CDS             7967. .9565
                     /gene="XF1854"
                     /note="hypothetical protein; identified by sequence
                     similarity; putative; ORF located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF84660.1"
                     /db_xref="GI:9106940"
```

```
  Query Match             2.1%;  Score 67;  DB 1;  Length 11476;
  Best Local Similarity   55.9%;  Pred. No. 0.00018;
  Matches  127;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy     1224 CGGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATA 1283
             ||| || |||||||| || ||   | ||| | ||   | |||||  |   ||   | ||  ||
Db     4978 CGGCACGTCCTTTTCAGCGCCTTTGGTATCTGGTGCAGCGGCGTTAGTGTGGGAGGTATT 5037

Qy     1284 CCCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGG 1343
              || |   | |  |||||     |||||   ||  ||||    |||||     ||  | ||
Db     5038 TCCATACTTTAATAACGATCTAGTGCGTCAGACATTGCTTGGGACGGCCACCGATCTTGG 5097

Qy     1344 TGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAA 1403
             |||    |||||||||       ||||| |  ||     ||| ||| ||| || ||  ||
Db     5098 TGCTCCTGGCGTGGACGAAGTATTCGGGTATGGCGCTCTGAATGTGGGCAAAGCGGTGCG 5157

Qy     1404 CGGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCGATACGAAAG 1450
              || || ||     |||   |  |||||  | |  | | |     || |
Db     5158 TGGGCCTGCCAAGTTTGATTGGGGCGATGTCAAGGTCAACTTCAATG 5204


RESULT 39
AF216702
LOCUS       AF216702                5423 bp    DNA     linear   BCT 06-MAR-2001
DEFINITION  Pseudomonas fluorescens PrtB (prtB) and lipase (lipA) genes,
            complete cds.
ACCESSION   AF216702
VERSION     AF216702.1  GI:8895499
KEYWORDS    .
SOURCE      Pseudomonas fluorescens.
  ORGANISM  Pseudomonas fluorescens
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 5423)
  AUTHORS   Woods,R.G., Burger,M., Beven,C.A. and Beacham,I.R.
  TITLE     The aprX-lipA operon of Pseudomonas fluorescens B52: a molecular
            analysis of metalloprotease and lipase production
  JOURNAL   Microbiology 147 (Pt 2), 345-354 (2001)
  MEDLINE   21097273
   PUBMED   11158351
REFERENCE   2  (bases 1 to 5423)
  AUTHORS   Woods,R.G., Burger,M., Beven,C.-A. and Beacham,I.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1999) Health Science, Griffith University,
            Parklands Drive, Goldcoast, Qld 4215, Australia
FEATURES             Location/Qualifiers
     source          1..5423
                     /organism="Pseudomonas fluorescens"
                     /strain="B52"
                     /db_xref="taxon:294"
     gene            575..3685
                     /gene="prtB"
     CDS             575..3685
                     /gene="prtB"
                     /note="similar to serine protease"
                     /codon_start=1
                     /transl_table=11
                     /product="PrtB"
                     /protein_id="AAF80995.1"
                     /db_xref="GI:8895500"
                     /translation="MNKRKSGLITATHTGPGYPLKALSCVPYGALLCCLASLGTQAA
                     PYVEAGKLGDAASWRSNEFKADCGLGAVHADAAYAAGYTGKGVKVGIFDQPVYAQHPE
                     FDSPGKVVTIVTEGNRQYTDPYIPVKAGDAFRYDGTPSKDSNGKLGNHGTHVGGIAAG
                     NRDGGPMHGVAFNAQILSAENGDPGPEDGIILGNDGAVYKAGWDGLVANGARIINNSW
                     GIGIGDQYAKGGRDPAFPNFTVNEAQAQFNVIRPILGTLAGGAYQGAIDAARSGVLTI
                     FAAGNDYNRNNPDAMSGLAYFVPDIAPNWLSVAALQQNPDTASADPYVISTFSSRCGY
                     AASFCVSAPGTKIFSSIINGTDLSNLTTDWANKNGTSMAAPHVAGAAAVLMERFPYMN
                     GDQVSTVLKTTATELRAPGIDALYGWGMINLGKAGNGPGMFITGADIPAEFRIDGAYG
                     NGQFVANLPGIGAVVDAGKPTQRLCTDVHCGLDVWSNNISGHGGLTKQGIGALVLTGN
                     NTYSGPTLVNQGLLAINGSVTSDVTVSQNGVVGGSGRIGSLSAKSGGTVAPGNSIGTL
                     NVAGNVTFDAGSTYAVELSPNSSDRIVPGGTATLMRTTVTLLLENSRTLLTNAQAESL
                     IGRQYNILQAAGGVTGDFGAVLPNYLFIGGTLNYAANGVQLDVARNANSFASVAVTPN
                     QRAVAAAADQLGAGNAVYESLLLAPTAASAQGAFQQLSGEIYPALQSALVNDSRYVRE
                     AVGERLRNGEMGATSETLDSRGNVWVKALGAWGKTDSRSDTAGYTTSLGGMLVGVDGA
                     LDEATRIGLVAGYSDTSLNMGNDTHSRASVDSYHFGAYAGHEIGALRLSGGATYSWHR
                     ADVKRDLQYGDVAGKQKVKVDASSTQVFGEAAYRVNLQPLALEPFANLAYVHLDTDGF
                     TEKGDAAALKGHDDSRDVVLSTLGMRALKTFNVNDHQQLEVSGTLGWQHNLSSTDSEQ
                     HLAFASGSPSFAVESAPMVRDAALVGARVSLALSKEARVNFDYNGLLASKDKVHGVGV
                     SLDWAF"
     gene            3814..5244
                     /gene="lipA"
     CDS             3814..5244
                     /gene="lipA"
                     /codon_start=1
                     /transl_table=11
                     /product="lipase"
                     /protein_id="AAF80996.1"
                     /db_xref="GI:8895501"
                     /translation="MGIFDYKNLGTEGSKTLFADAMAITLYSYHNLDNGFAVGYQHNG
                     LGLGLPATLVGALLGSTDSQGVIPGIPWNPDSEKAALEAVQKAGWTPISASALGYAGK
                     VDARGTFFGEKAGYTTAQVEVLGKYDDAGKLLEIGIGFRGTSGPRETLISDSIGDLIS
                     DLLAALGPKDYAKNYAGEAFGGLLKNVADYAGAHGLTGKDVVVSGHSLGGLAVNSMAD
                     LSNYKWAGFYKDANYVAYASPTQSAGDKVLNIGYENDPVFRALDGSSFNLSSLGVHDK
                     PHESTTDNIVSFNDHYASTLWNVLPFSIVNLPTWVSHLPTAYGDGMTRILESGFYDQM
                     TRDSTVIVANLSDPARANTWVQDLNRNAEPHKGNTFIIGSDGNDLIQGGNGADFIEGG
                     KGNDTIRDNSGHNTFLFSGHFGNDRVIGYQPTDKLVFKDVQGSTDLRDHAKVVGADTV
                     LTFGADSVTLVGVGHGGLWTEGVVIG"
BASE COUNT     1093 a   1659 c   1734 g    937 t
ORIGIN

  Query Match             2.0%;  Score 65;  DB 1;  Length 5423;
  Best Local Similarity   56.2%;  Pred. No. 0.0005;
  Matches  122;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;

Qy     1225 GGAACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATAC 1284
             || || ||| |  ||||||||  |||  ||||| | |||||  || || || |     | |
Db     1679 GGCACTTCCATGGCCGCACCCCACGTGGCCGGCGCTGCGGCGGTGTTGATGGAACGCTTC 1738

Qy     1285 CCGTGGATGAGCAACGACAACCTGCGTACCACGCTGCTGACAACGGCTCAGGACATCGGT 1344
             ||||  |||| |  |||  |  |      |||  |||     || || ||     ||  || |
Db     1739 CCGTACATGAACGGCGATCAGGTCTCCACCGTGCTCAAAACCACCGCCACCGAACTCCGC 1798

Qy     1345 GCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAAC 1404
             ||     ||  |  |||        || |||||||||  |  |   | ||| ||||||   |||
Db     1799 GCGCCAGGGATCGACGCGCTGTACGGCTGGGGCATGATCAACCTGGGCAAGGCCGGCAAC 1858

Qy     1405 GGACCCGCGTCCTTTCCGTTCGGCGACTTTACCGCCG 1441
             || || |       ||        |||||    || | | |
Db     1859 GGCCCGGGAATGTTCATTACCGGCGCTGATATCCCGG 1895


RESULT 40
E32986
LOCUS       E32986                 10732 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Gene encoding cellulose synthesizer.
ACCESSION   E32986
VERSION     E32986.1  GI:13022340
KEYWORDS    JP 2000060568-A/1.
SOURCE      Vigna angularis.
  ORGANISM  Vigna angularis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Vigna.
REFERENCE   1  (bases 1 to 10732)
  AUTHORS   Koichi,M., Tomohiko,K., Shigeru,S. and Daisuke,S.
  TITLE     Gene encoding cellulose synthesizer
  JOURNAL   Patent: JP 2000060568-A 1 29-FEB-2000;
            KOICHI MIZUNO,MITSUI GIYOUSAI SHOKUBUTSU BIO KENKYUSHO
COMMENT     OS   Vigna angularis
            PN   JP 2000060568-A/1
            PD   29-FEB-2000
            PF   26-AUG-1998 JP 1998239998
            PR
            PI   KOICHI MIZUNO,TOMOHIKO KATO,SHIGERU SATO,DAISUKE SHIBATA PC
            C12N15/09,A01H5/00,C07K16/40,C12N5/10,C12N9/10,C12P21/02// PC
```

```
            (C12N15/09,C12R1:91),(C12N5/10,C12R1:91),(C12P21/02,C12R1:91), PC
            C12N15/00,
            PC   C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,C12R1:91) CC
            FH   Key              Location/Qualifiers
            FT   CDS              (1). .(3375).
FEATURES             Location/Qualifiers
     source          1. .10732
                     /organism="Vigna angularis"
                     /db_xref="taxon:3914"
BASE COUNT     3149 a   1212 c   2074 g   2046 t   2251 others
ORIGIN

  Query Match             1.9%;  Score 62;  DB 6;  Length 10732;
  Best Local Similarity   13.9%;  Pred. No. 0.0026;
  Matches   98;  Conservative  300;  Mismatches  304;  Indels    3;  Gaps    2;

Qy     2388 CCACAACGCTACGGGTCTGCGCGTCATCGCGCAAACCCAACAGGACGGTGGAACGTGGGA 2447
            :: :||: : | ||: :||:: : : |:::: :: :  :| :|:|  : | :  : || :
Db     9208 HSGYAAMTSRARGGSRMTGYSASASNTHRSRTHRTYRGVAARGVAGYSGGARGASNGGAR 9267

Qy     2448 ACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCCGCGAA 2507
             :::  |  ||::|||:  :: ::    |:: ::::: ||::| | :    |  |: |:
Db     9268 GGHSRARGARGTHRAAGVASRHSYSGNGNGSRTHRYSYSAAYSGGCYSGNARGARGVAAR 9327

Qy     2508 AACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACACAGCACATGGAGCGA 2567
             : :| | :|| |   ::: :  : |:  :  : :|   : ||   :|::: : :: :
Db     9328 GYSARGRGYRAAAAAAYSMTYSASNGVAGMTGYARGASSRHGARGARGSRYSSRTHRASN 9387

Qy     2568 AAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGC 2627
             ::  : : :||  :: ::  ::: :::       : |      | :  :   | : ::: :
Db     9388 SRVAGSRSRVAASTHRSRRGTHRRTHRYSASNTHRTHRGGNCYSAAVAGGGASNYSVAYS 9447

Qy     2628 GGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAG 2687
            |:: :  : :: :    :::|:|: :| :      ::  ::::: ::   |::::  :::
Db     9448 GSRASNARGYSASNGGNHSRARGVAMTTHRARGTHRAASRYSHSGSRGNGASHSYSGYHS 9507

Qy     2688 CCGCAGCACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCT 2747
            : :   :::::   :  |::: : ::   :::     |:::|: ::      : :: ||:
Db     9508 SRVAGHSHSSRASNVAAASRHSGHSRHAASRSRASAAGYSRASASYSTHRSRCYSAAGSR 9567

Qy     2748 GGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTCGAAGG 2807
             :   : | :|   :::: : :: :::::  ::  :  ::::||    :|: : : |  :
Db     9568 TRAASRAASRGA--SHSHARGSRGYSYSYSGRSRCYSYSSRVAGYAASRASASNMTASAS 9625

Qy     2808 CGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGG 2867
            ::  :||::  : :|:     ::|| ::|    :::  :::: |:| ::|     ::  | :|
Db     9626 HVAGMTGYSAAVAVASRVAGYSAATHRGNSRRSRAASRYSAAVATHRGASNSRSRGTHRG 9685

Qy     2868 CTGGAGCGGCAACAGCCTCACTGAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTC 2927
            :| : : :: ::  | :::     :|: :     | :|::| :     :|: :    :    |:
Db     9686 YTHRYSGSRTHRRGGGYYSGRVARASHSSRGNTHRTHRGTHRTHRRGVAVAASMTGRMTS 9745

Qy     2928 GCAACCCTTGAGCGATAAAGCCGTCCTGTTTGCAACGGCGGGCGTGGAACGCGACCTGAA 2987
             :  :::    :: :::: |     :: :: :  :  :     :: :|: |:| |:: :|::| :
Db     9746 RASHSSRASSRYSSRYSASNTHRCYSSRASHYSGYTHRVARGY-TRGNASVAVAYSTHRG 9804

Qy     2988 CGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCGACTGCAGCAACCGGCAAGAC 3047
                ::| ::: ::|::|: : : ::    |   :|::: |:   :: ::|:: ::: :|
Db     9805 GNASNHSVATHRHSYSGYRASASHSASYSVAAAARGHSVAASNASHASHASHASSRSRAR 9864

Qy     3048 GGGGGCACGCAATATGCCGCACACCCGCCTGGTTGCCGGTCTGGG 3092
            || :::|::    ::    :    :   |:: :|:::    :::  : :::
Db     9865 GGYSRVAHSASTHRGNASNRRGNCYSHSCYSSRCYSSRASSRSRV 9909

Search completed: January 27, 2003, 19:00:20
Job time : 7910 secs
```

Baskar
09/830433


```
    FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,
    JICST-EPLUS, JAPIO' ENTERED AT 10:13:04 ON 28 JAN 2003
L1          105 SEA ABB=ON  PLU=ON  "AUJAME L"?/AU
L2           29 SEA ABB=ON  PLU=ON  "BOUCHARDON A"?/AU
L3           79 SEA ABB=ON  PLU=ON  ("RENAULD MONGENIE G"? OR "MONGENIE
                RENAULD G"? OR "MONGENIE G"? OR "RENAULD G"?)/AU
L4           65 SEA ABB=ON  PLU=ON  "ROKBI B"?/AU
L5          309 SEA ABB=ON  PLU=ON  "NASSIF X"?/AU
L6          208 SEA ABB=ON  PLU=ON  "TINSLEY C"?/AU
L7         1372 SEA ABB=ON  PLU=ON  "PERRIN A"?/AU
L8            2 SEA ABB=ON  PLU=ON  L1 AND L2 AND L3 AND L4 AND L5 AND
                L6 AND L7
L9            5 SEA ABB=ON  PLU=ON  L1 AND (L2 OR L3 OR L4 OR L5 OR L6
                OR L7)
L10           2 SEA ABB=ON  PLU=ON  L2 AND (L3 OR L4 OR L5 OR L6 OR L7)

L11           8 SEA ABB=ON  PLU=ON  L3 AND (L4 OR L5 OR L6 OR L7)
L12           2 SEA ABB=ON  PLU=ON  L4 AND (L5 OR L6 OR L7)
L13          52 SEA ABB=ON  PLU=ON  L5 AND (L6 OR L7)
L14          19 SEA ABB=ON  PLU=ON  L6 AND L7
L15        2079 SEA ABB=ON  PLU=ON  L1 OR L2 OR L3 OR L4 OR L5 OR L6 OR
                L7
L16         293 SEA ABB=ON  PLU=ON  (L13 OR L15) AND NEISSER?
L17          66 SEA ABB=ON  PLU=ON  L16 AND (VACCIN? OR IMMUNIS? OR
                IMMUNIZ?)
L18          26 SEA ABB=ON  PLU=ON  L17 AND PATHOGEN?
L19          52 SEA ABB=ON  PLU=ON  L8 OR L9 OR L10 OR L11 OR L12 OR L14
                OR L18
L20          25 DUP REM L19 (27 DUPLICATES REMOVED)
```

— Author(s)

L20 ANSWER 1 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 1

ACCESSION NUMBER: 2002:908217 HCAPLUS

TITLE: Comparative genomics identifies the genetic islands that distinguish Neisseria meningitidis, the agent of cerebrospinal meningitis, from other Neisseria species

AUTHOR(S): **Perrin, Agnes**; Bonacorsi, Stephane; Carbonnelle, Etienne; Talibi, Driss; Dessen, Philippe; Nassif, Xavier; **Tinsley, Colin**

CORPORATE SOURCE: INSERM U5701 Faculte de Medecine Necker, Paris, 75015, Fr.

SOURCE: Infection and Immunity (2002), 70(12), 7063-7072 CODEN: INFIBR; ISSN: 0019-9567

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Neisseria meningitidis colonizes the nasopharynx and, unlike commensal Neisseria species, is capable of entering the bloodstream, crossing the blood-brain barrier, and invading the meninges. The other pathogenic Neisseria species, Neisseria gonorrhoeae, generally causes an infection which is localized to the genitourinary tract. In order to investigate the genetic basis of this difference in disease profiles, we used a strategy of genomic comparison. We used DNA arrays to compare the genome of N. meningitidis with those of N. gonorrhoeae and Neisseria lactamica, a commensal of the nasopharynx. We thus identified sequences conserved among a representative set of virulent strains which are either specific to N. meningitidis or shared with N. gonorrhoeae but absent from N. lactamica. Though

these bacteria express dramatically different pathogenicities, these meningococcal sequences were limited and, in contrast to what has been found in other pathogenic bacterial species, they are not organized in large chromosomal islands.

REFERENCE COUNT: 39 THERE ARE 39 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 2 OF 25 MEDLINE
ACCESSION NUMBER: 2002158631 MEDLINE
DOCUMENT NUMBER: 21887396 PubMed ID: 11890539
TITLE: Genomics of **Neisseria** meningitidis.
AUTHOR: **Nassif Xavier**
CORPORATE SOURCE: INSERM U411, Faculte de Medecine Necker-Enfants Malades, Universite Rene Descartes, Paris, France.. nassif@necker.fr
SOURCE: INTERNATIONAL JOURNAL OF MEDICAL MICROBIOLOGY, (2002 Feb) 291 (6-7) 419-23. Ref: 28
Journal code: 100898849. ISSN: 1438-4221.
PUB. COUNTRY: Germany: Germany, Federal Republic of
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200209
ENTRY DATE: Entered STN: 20020314
Last Updated on STN: 20020928
Entered Medline: 20020927

AB An important feature of disease caused by **Neisseria** meningitidis is the propensity to invade the meninges. Much progress has been made in our understanding of how this **pathogen** circumvents the physical properties of this cellular barrier. This review will address the new possibilities offered by the recent availability of meningococcal genome sequences.

L20 ANSWER 3 OF 25 MEDLINE DUPLICATE 2
ACCESSION NUMBER: 2002712859 MEDLINE
DOCUMENT NUMBER: 22362872 PubMed ID: 12474388
TITLE: Neisseria microarrays.
AUTHOR: **Tinsley Colin R; Perrin Agnes;**
Borezee Elise; Nassif Xavier
CORPORATE SOURCE: INSERM U570, Faculte de Medecine Necker-Enfants Malades, 75730 Paris, France.
SOURCE: METHODS IN ENZYMOLOGY, (2002) 358 188-207.
Journal code: 0212271. ISSN: 0076-6879.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200301
ENTRY DATE: Entered STN: 20021217
Last Updated on STN: 20030122
Entered Medline: 20030121

L20 ANSWER 4 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2001:477334 BIOSIS
DOCUMENT NUMBER: PREV200100477334

TITLE: Hemoglobin receptors from **neisseriae.**
AUTHOR(S): Stojiljkovic, Igor; So, Magdalene; Hwa, Vivian; Heffron, Fred; **Nassif, Xavier (1)**
CORPORATE SOURCE: (1) Paris France
ASSIGNEE: Oregon Health Sciences University
PATENT INFORMATION: US 6277382 August 21, 2001
SOURCE: Official Gazette of the United States Patent and Trademark Office Patents, (Aug. 21, 2001) Vol. 1249, No. 3, pp. No Pagination. e-file.
ISSN: 0098-1133.
DOCUMENT TYPE: Patent
LANGUAGE: English

AB The present invention relates to novel bacterial hemoglobin receptor proteins and genes that encode such proteins. The invention is directed toward the isolation, characterization, diagnostic and therapeutic use of bacterial hemoglobin receptor proteins, nucleic acids encoding such proteins, recombinant expression constructs comprising such nucleic acids and cells transformed therewith, and antibodies and epitopes of such hemoglobin receptor proteins. The invention relates particularly to hemoglobin receptor proteins and genes encoding such proteins from **Neisseria** species, especially N. meningitidis and serotypes thereof, and N. gonorrhoeae. Methods for the diagnostic and therapeutic use of the proteins, epitopes, antibodies and nucleic acids of the invention are also provided, including the use of proteins, epitopes, antibodies and nucleic acids of the invention for the production of **vaccines** effective in providing **immunization** of human against infection by **pathogenic** bacteria of **Neisseria** species.

L20 ANSWER 5 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2001:334753 BIOSIS
DOCUMENT NUMBER: PREV200100334753
TITLE: Preparation of a multicombinatorial library of antibody gene expression vectors.
AUTHOR(S): Sodoyer, Regis (1); **Aujame, Luc**; Geoffroy, Frederique; **Bouchardon, Annabelle**
CORPORATE SOURCE: (1) Saint Foy les Lyon France
ASSIGNEE: Pasteur Merieux Serums & Vaccins, Lyon Cedex, France
PATENT INFORMATION: US 6174708 January 16, 2001
SOURCE: Official Gazette of the United States Patent and Trademark Office Patents, (Jan. 16, 2001) Vol. 1242, No. 3, pp. No Pagination. e-file.
ISSN: 0098-1133.
DOCUMENT TYPE: Patent
LANGUAGE: English

AB On the basis of a first repertoire of genes coding for a population of one of two kinds of polypeptides capable of being optionally covalently combined, particularly variable regions of either the antibody light chain type or the antibody heavy chain type, and at least one gene coding for the other type of polypeptide, particularly a variable region of the other type, an antibody chain or preferably a second repertoire of genes coding for a population of said other type, the genes from the first repertoire are inserted into a first vector to form a population of vectors carrying the various genes of said first repertoire, and said gene of said other type or the genes from said second repertoire is/are inserted into a

THIS PAGE BLANK (USPTO)

second vector. Both starting vectors have means enabling each to exchange one part by one or more irreversible recombinations to generate recombinant final vectors of which one contains a gene of one of said types and a gene of the other type.

L20 ANSWER 6 OF 25 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2001-082916 [10] WPIDS
DOC. NO. NON-CPI: N2001-063334
DOC. NO. CPI: C2001-024200
TITLE: Immunogenic polypeptides derived from **Neisseria** meningitidis and the nucleic acids that encode them, useful for diagnosing and **vaccinating** against **Neisseria** infections e.g. bacteremia and meningitis.
DERWENT CLASS: B04 D16 S03
INVENTOR(S): **NASSIF, X**; **TINSLEY, C**; ACHTMAN, M; KLEE, S; MERKER, P
PATENT ASSIGNEE(S): (INRM) INSERM INST NAT SANTE & RECH MEDICALE; (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN
COUNTRY COUNT: 95
PATENT INFORMATION:

```
PATENT NO    KIND DATE     WEEK      LA   PG
---------------------------------------------
EP 1069133     A1 20010117 (200110)* EN  232
    R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK
       NL PT RO SE SI
WO 2001004150 A2 20010118 (200110)   EN
   RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC
       MW MZ NL OA PT SD SE SL SZ TZ UG ZW
    W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE
       DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG
       KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ
       PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN
       YU ZA ZW
AU 2000068254 A  20010130 (200127)
EP 1194446     A2 20020410 (200232)  EN
    R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK
       NL PT RO SE SI
```

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
EP 1069133	A1	EP 1999-401764	19990713
WO 2001004150	A2	WO 2000-EP6943	20000705
AU 2000068254	A	AU 2000-68254	20000705
EP 1194446	A2	EP 2000-956222	20000705
		WO 2000-EP6943	20000705

FILING DETAILS:

PATENT NO	KIND		PATENT NO
AU 2000068254	A	Based on	WO 200104150
EP 1194446	A2	Based on	WO 200104150

PRIORITY APPLN. INFO: EP 1999-401764 19990713

AN 2001-082916 [10] WPIDS
AB EP 1069133 A UPAB: 20010220
NOVELTY - Immunologically active polypeptides (I) derived from the Gram negative bacteria **Neisseria** meningitidis, and the nucleic acids (II) that encode them, are new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following:

(1) an isolated polypeptide (I) comprising an amino acid sequence that has at least 70% identity to 44 defined amino acid sequences ((A1)-(A44)) given in the specification;

(2) an immunogenic fragment of (I) which comprises (A1)-(A44);

(3) an isolated polynucleotide (II) comprising a nucleotide sequence encoding (I) (which has at least 70% to (A1)-(A44) over its entire length), or a sequence complementary to (II);

(4) an expression vector (III) or a recombinant live microorganism comprising (II);

(5) a host cell (IV) comprising (III), or a membrane of (IV), that expresses a polypeptide comprising an amino acid sequence with at least 70% identity to (A1)-(A44);

(6) a process (V) for producing a polypeptide comprising an amino acid sequence with at least 70% identity to (A1)-(A44), comprising culturing the host cell (IV) under suitable conditions for expression of the polypeptide and recovering the polypeptide from the culture medium;

(7) a process (VI) for expressing the polynucleotide (II), comprising transforming a host cell with an expression vector comprising (II) and culturing the host cell under conditions suitable for expression of the polypeptide;

(8) **vaccine** compositions (VII) comprising (I) and/or (II);

(9) antibody (VIII) immuno-specific for (I); and

(10) a method for diagnosing a **Neisseria** infection, comprising identifying (I) or (VIII) in a sample from the subject animal.

ACTIVITY - Antibacterial.

MECHANISM OF ACTION - **Vaccine.**

Rabbit antiserum produced in response to **vaccination** with the polypeptides killed 65% of parenterally administered meningococcus (strain 8013) with in 20 minutes of contact and all of the bacteria within 60 minutes. Pre-immune serum (taken prior to **immunization**) was found to have killed no bacteria after 20 minutes and only half after 60 minutes.

USE - The nucleic acids and the polypeptides they encode may be used to **vaccinate** subjects against infection by **Neisseria** meningitidis bacteria according to standard methodologies. The antibodies produced in response to the polypeptides and/or polynucleotides may also be used to treat N. meningitidis infections or as diagnostic reagents in immunoassays to detect infections (claimed). N. meningitidis is a **pathogen** involved in, for example, bacteremia and meningitis.
Dwg.0/50

L20 ANSWER 7 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2001:443973 BIOSIS
DOCUMENT NUMBER: PREV200100443973
TITLE: Production of **Neisseria** meningitidis transferrin-binding protein B by recombinant Bordetella pertussis.

AUTHOR(S): Coppens, Isabelle; Alonso, Sylvie; Antoine, Rudy; Jacob-Dubuisson, Francoise; **Renauld-Mongenie, Genevieve**; Jacobs, Eric; Locht, Camille (1)
CORPORATE SOURCE: (1) Laboratoire de Microbiologie Genetique et Moleculaire, INSERM U447, Institut Pasteur de Lille, 1 Rue du Prof. Calmette, F-59019, Lille Cedex: camille.locht@pasteur-lille.fr France
SOURCE: Infection and Immunity, (September, 2001) Vol. 69, No. 9, pp. 5440-5446. print.
ISSN: 0019-9567.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English
AB **Neisseria** meningitidis serogroup B infections are among the major causes of fulminant septicemia and meningitis, especially severe in young children, and no broad **vaccine** is available yet. Because of poor immunogenicity of the serogroup B capsule, many efforts are now devoted to the identification of protective protein antigens. Among those are PorA and, more recently, transferrin-binding protein B (TbpB). In this study, TbpB of N. meningitidis was genetically fused to the N-terminal domain of the Bordetella pertussis filamentous hemagglutinin (FHA), and the fha-tbpB hybrid gene was expressed in B. pertussis either as a plasmid-borne gene or as a single copy inserted into the chromosome. The hybrid protein was efficiently secreted by the recombinant strains, despite its large size, and was recognized by both anti-FHA and anti-TbpB antibodies. A single intranasal administration of recombinant virulent or pertussis-toxin-deficient, attenuated B. pertussis to mice resulted in the production of antigen-specific systemic immunoglobulin G (IgG), as well as local IgG and IgA. The anti-TbpB serum antibodies were of the IgG1, IgG2a, and IgG2b isotypes and were found to express complement-mediated bactericidal activity against N. meningitidis. These observations indicate that recombinant B. pertussis may be a promising vector for the development of a mucosal **vaccine** against serogroup B meningococci.

L20 ANSWER 8 OF 25 MEDLINE DUPLICATE 3
ACCESSION NUMBER: 2001232892 MEDLINE
DOCUMENT NUMBER: 21112216 PubMed ID: 11173033
TITLE: Meningococcal **pathogenesis**: at the boundary between the pre- and post-genomic eras.
AUTHOR: **Tinsley C; Nassif X**
CORPORATE SOURCE: INSERM U411, Faculte de Medecine Necker - Enfants Malades, Universite Rene Descartes, 75730 Cedex 15, Paris, France.
SOURCE: CURRENT OPINION IN MICROBIOLOGY, (2001 Feb) 4 (1) 47-52. Ref: 49
Journal code: 9815056. ISSN: 1369-5274.
PUB. COUNTRY: England: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200105
ENTRY DATE: Entered STN: 20010517
Last Updated on STN: 20010517

Entered Medline: 20010503

AB Meningococcal disease remains an important public health burden worldwide and, indeed, cause of death, particularly in poorer countries. The rapidly progressive nature of infections means that antibiotic therapy often comes too late. **Vaccines** are of limited efficacy in infants, one of the most vulnerable age groups, and do not exist for bacteria of serogroup B. Hence, much remains to be achieved in terms of **vaccine** design and the understanding of the **pathogenesis** of meningococcal disease. The causative bacterium, **Neisseria** meningitidis, is usually a commensal of the nasopharynx. Factors that lead to the invasion of the bloodstream, often followed by the crossing of the blood-brain barrier and meningitis, may be partly host- and partly bacterium-dependent, but are ill-understood. It is hoped that, taken together with the fundamental knowledge gained from biochemical and genetic studies, the huge amount of new information made available with the recent publication of the genome sequences will help to unlock more of the secrets of the lifestyle and **pathogenic** potential of this still poorly understood **pathogen**.

L20 ANSWER 9 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2002:211711 BIOSIS
DOCUMENT NUMBER: PREV200200211711
TITLE: New pathogenicity factors from genomic comparison.
AUTHOR(S): **Tinsley, C. R. (1); Perrin, A. (1)**
; Nassif, X. (1)
CORPORATE SOURCE: (1) Faculte de Medecine Necker, INSERM U411, 156 Rue de Vaugirard, 75015, Paris France
SOURCE: International Journal of STD & AIDS, (2001) Vol. 12, No. Supplement 2, pp. 22-23. print.
Meeting Info.: International Congress of Sexually Transmitted Infections Berlin, Germany June 24-27, 2001 International Union Against Sexually Transmitted Infections
. ISSN: 0956-4624.
DOCUMENT TYPE: Conference
LANGUAGE: English

L20 ANSWER 10 OF 25 MEDLINE
ACCESSION NUMBER: 2000210556 MEDLINE
DOCUMENT NUMBER: 20210556 PubMed ID: 10755929
TITLE: Microbiology. A furtive **pathogen** revealed.
COMMENT: Comment on: Science. 2000 Mar 10;287(5459):1809-15
Comment on: Science. 2000 Mar 10;287(5459):1816-24
AUTHOR: **Nassif X**
CORPORATE SOURCE: INSERM Unit 411, Facult| de M|decine Necker-Enfants Malades, 156 Rue de Vaugirard, 75015 Paris, France.. nassif@necker.fr
SOURCE: SCIENCE, (2000 Mar 10) 287 (5459) 1767-8.
Journal code: 0404511. ISSN: 0036-8075.
PUB. COUNTRY: United States
DOCUMENT TYPE: Commentary
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200004
ENTRY DATE: Entered STN: 20000413
Last Updated on STN: 20000413

Entered Medline: 20000403

L20 ANSWER 11 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 4
ACCESSION NUMBER: 2000:314839 HCAPLUS
DOCUMENT NUMBER: 132:330635
TITLE: Genes and proteins specific for Neisseria meningitidis and their use in vaccination
INVENTOR(S): **Aujame, Luc; Bouchardon, Annabelle; Renauld-Mongenie, Genevieve; Rokbi, Bachra; Nassif, Xavier; Tinsley, Colin ; Perrin, Agnes**
PATENT ASSIGNEE(S): Pasteur Merieux Serums et Vaccins, Fr.; Institut National de la Sante et de la Recherche Medicale (INSERM)
SOURCE: PCT Int. Appl., 187 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: French
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000026375	A2	20000511	WO 1999-FR2643	19991028
WO 2000026375	A3	20000817		

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
FR 2785293	A1	20000505	FR 1998-13693	19981030
FR 2785293	B1	20020705		
AU 9963479	A1	20000522	AU 1999-63479	19991028
EP 1129195	A2	20010905	EP 1999-950875	19991028

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO

PRIORITY APPLN. INFO.: FR 1998-13693 A 19981030
WO 1999-FR2643 W 19991028

AB The invention concerns nucleic acids coding for polypeptides specific for Neisseria meningitidis, the corresponding polypeptides, and their diagnostic and therapeutic applications. Thus, genes and proteins found in N. meningitidis but not in N. lactamica were identified and sequenced.

L20 ANSWER 12 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 5
ACCESSION NUMBER: 2000:658423 HCAPLUS
DOCUMENT NUMBER: 133:248066
TITLE: Sequence and recombinant production of **Neisseria** meningitidis serotype B gene hmbR hemoglobin receptor
INVENTOR(S): Stojiljkovic, Igor; So, Magdalene; Hwa, Vivian; Heffron, Fred; **Nassif, Xavier**
PATENT ASSIGNEE(S): USA

SOURCE: U.S., 87 pp., Cont.-in-part of U.S. 5,698,438.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6121037	A	20000919	US 1995-537361	19951002
US 5698438	A	19971216	US 1994-326670	19941018
CA 2203116	AA	19960425	CA 1995-2203116	19951017
WO 9612020	A2	19960425	WO 1995-US13623	19951017
WO 9612020	A3	19960523		

W: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT

RW: KE, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
HU 77048	A2	19980302	HU 1997-1963	19951017
JP 10508469	T2	19980825	JP 1995-513482	19951017
NO 9701768	A	19970603	NO 1997-1768	19970417
FI 9701634	A	19970616	FI 1997-1634	19970417
US 6123942	A	20000926	US 1997-990470	19971215

PRIORITY APPLN. INFO.:
US 1994-326670 A2 19941018
US 1995-537361 A 19951002
WO 1995-US13623 W 19951017

AB The invention claims a DNA mol. encoding a Hb receptor protein from **Neisseria** meningitidis serotype B. The invention also claims the use of said DNA in construction of an expression construct and cells transformed with said expression construct for the recombinant prodn. of Hb receptor. The invention further provides a nucleic acid mol. fully complementary to N. meningitis serotype B Hb receptor DNA as a diagnostic agent. The DNA sequence, as well as the corresponding amino acid sequence of N. meningitis serotype B Hb receptor are provided. In the example section, the invention provided methods used clone Hb receptor genes (hmbR) from N. meningitis serotypes A and C, as well as from N. gonorrhoeae. The exampled also provided the DNA sequences, as as corresponding amino acid sequences, of these cloned Hb receptor. The invention further showed that: (1) the hmbR gene product is a Hb receptor in N. meningitis; (2) N. meningitis Hb receptor of the invention is TonB-dependent and (3) an IS1106 element is located downstream of hmbR gene in N. meningitis serotype C. Still further the invention discussed use of the proteins, epitopes, antibodies and nucleic acids of the invention for the prodn. of **vaccines** effective in providing **immunization** of a human against infection by **pathogenic** bacteria of **Neisseria** species.

REFERENCE COUNT: 49 THERE ARE 49 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 13 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2001:231857 BIOSIS
DOCUMENT NUMBER: PREV200100231857

TITLE: Bacterial hemoglobin receptor gene.
AUTHOR(S): Stojiljkovic, Igor (1); So, Magdalene; Hwa, Vivian; Heffron, Fred; **Nassif, Xavier**
CORPORATE SOURCE: (1) Portland, OR USA
ASSIGNEE: Oregon Health Sciences University
PATENT INFORMATION: US 6123942 September 26, 2000
SOURCE: Official Gazette of the United States Patent and Trademark Office Patents, (Sep. 26, 2000) Vol. 1238, No. 4, pp. No Pagination. e-file.
ISSN: 0098-1133.
DOCUMENT TYPE: Patent
LANGUAGE: English

AB The present invention relates to a novel bacterial hemoglobin receptor protein and genes that encode such a protein. The invention is directed toward the isolation, characterization, diagnostic and therapeutic use of a bacterial hemoglobin receptor protein, nucleic acid encoding such a protein, recombinant expression constructs comprising such nucleic acids and cells transformed therewith, and antibodies and epitopes of such hemoglobin receptor proteins. The invention relates particularly to hemoglobin receptor proteins and genes encoding such proteins from **Neisseria** species, especially N. meningitidis. Methods for the diagnostic and therapeutic use of the proteins, epitopes, antibodies and nucleic acids of the invention are also provided, including the use of the proteins, epitopes, antibodies and nucleic acids of the invention for the production of **vaccines** effectinve in providing **immunization** of a human against infection by **pathogenic** bacteria of **Neisseria** species.

L20 ANSWER 14 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 6
ACCESSION NUMBER: 2000:603693 HCAPLUS
DOCUMENT NUMBER: 134:52089
TITLE: Allelic diversity of the two transferrin binding protein B gene isotypes among a collection of Neisseria meningitidis strains representative of serogroup B disease: implication for the composition of a recombinant TbpB-based vaccine
AUTHOR(S): **Rokbi, Bachra**; **Renauld-Mongenie, Genevieve**; Mignon, Michele; Danve, B.; Poncet, David; Chabanel, Christophe; Caugant, Dominique A.; Quentin-Millet, Marie-Jose
CORPORATE SOURCE: Aventis Pasteur, Marcy-L'Etoile, 69280, Fr.
SOURCE: Infection and Immunity (2000), 68(9), 4938-4947
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The distribution of the two isotypes of tbpB in a collection of 108 serogroup B meningococcal strains belonging to the four major clonal groups assocd. with epidemic and hyperendemic disease (the ET-37 complex, the ET-5 complex, lineage III, and cluster A4) was detd. Isotype I strains (with a 1.8-kb tbpB gene) was less represented than isotype II strains (19.4 vs. 80.6%). Isotype I was restricted to the ET-37 complex strains, while isotype II was found in all four clonal complexes. The extent of the allelic diversity of tbpB in these two groups was studied by PCR restriction anal. and sequencing of 10 new tbpB genes. Four major tbpB gene variants were characterized: B16B6 (representative of isotype I) and M982, BZ83,

and 8680 (representative of isotype II). The relevance of these variants was assessed at the antigenic level by the detn. of cross-bactericidal activity of purified IgG prepns. raised to the corresponding recombinant TbpB (rTbpB) protein against a panel of 27 strains (5 of isotype I and 22 of isotype II). The results indicated that rTbpB corresponding to each variant was able to induce cross-bactericidal antibodies. However, the no. of strains killed with an anti-rTbpB serum was slightly lower than that obtained with an anti-TbpA+B complex. None of the sera tested raised against an isotype I strain was able to kill an isotype II strain and vice versa. None of the specific antisera tested (anti-rTbpB or anti-TbpA+B complex) was able to kill all of the 22 isotype II strains tested. Moreover, using sera raised against the C-terminus domain of TbpB M982 (amino acids 352 to 691) or BZ83 (amino acids 329 to 669) fused to the maltose-binding protein, cross-bactericidal activity was detected against 12 and 7 isotype II strains, resp., of the 22 tested. These results suggest surface accessibility of the C-terminal end of TbpB. Altogether, these results show that although more than one rTbpB will be required in the compn. of a TbpB-based vaccine to achieve a fully cross-bactericidal activity, rTbpB and its C terminus were able by themselves to induce cross-bactericidal antibodies.

REFERENCE COUNT: 46 THERE ARE 46 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 15 OF 25 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 2000097426 EMBASE
TITLE: A furtive **pathogen** revealed.
AUTHOR: **Nassif X.**
CORPORATE SOURCE: X. Nassif, INSERM Unit 411, Fac. de Med. Necker-Enfants Naiades, 156 ~~Rue de Vaugirard,~~ 75015 Paris, France. nassif@~~necker~~.fr
SOURCE: Science, ~~(10 Mar 2000)~~ 287/5459 (1767-1768). ~~Refs: 7~~
ISSN: 0036-8075 CODEN: SCIEAS
COUNTRY: United States
DOCUMENT TYPE: Journal; (Short Survey)
FILE SEGMENT:
004 Microbiology
008 Neurology and Neurosurgery
017 Public Health, Social Medicine and Epidemiology
026 Immunology, Serology and Transplantation
037 Drug Literature Index
LANGUAGE: English

L20 ANSWER 16 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 7
ACCESSION NUMBER: 1999:709902 HCAPLUS
DOCUMENT NUMBER: 132:59905
TITLE: Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to the pathogenic Neisseria species
AUTHOR(S): **Perrin, Agnes**; Nassif, Xavier; **Tinsley, Colin**
CORPORATE SOURCE: Laboratoire de Microbiologie, INSERM U411, Faculte de Medecine Necker-Enfants Malades, Paris, 75015, Fr.

SOURCE: Infection and Immunity (1999), 67(11), 6119-6129
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Neisseria meningitidis and Neisseria gonorrhoeae give rise to dramatically different diseases. Their interactions with the host, however, do share common characteristics: they are both human pathogens which do not survive in the environment and which colonize and invade mucosa at their port of entry. It is therefore likely that they have common properties that might not be found in nonpathogenic bacteria belonging to the same genetically related group, such as Neisseria lactamica. Their common properties may be detd. by chromosomal regions found only in the pathogenic Neisseria species. To address this issue, we used a previously described technique (C. R. Tinsley and X. Nassif, Proc. Natl. Acad. Sci. USA 93:11109-11114, 1996) to identify sequences of DNA specific for pathogenic neisseriae and not found in N. lactamica. Sequences present in N. lactamica were phys. subtracted from the N. meningitidis Z2491 sequence and also from the N. gonorrhoeae FA1090 sequence. The clones obtained from each subtraction were tested by Southern blotting for their reactivity with the three species, and only those which reacted with both N. meningitidis and N. gonorrhoeae (i.e., not specific to either one of the pathogens) were further investigated. In a first step, these clones were mapped onto the chromosomes of both N. meningitidis and N. gonorrhoeae. The majority of the clones were arranged in clusters extending up to 10 kb, suggesting the presence of chromosomal regions common to N. meningitidis and N. gonorrhoeae which distinguish these pathogens from the commensal N. lactamica. The sequences surrounding these clones were detd. from the N. meningitidis genome-sequencing project. Several clones corresponded to previously described factors required for colonization and survival at the port of entry, such as IgA protease and PilC. Others were homologous to virulence-assocd. proteins in other bacteria, demonstrating that the subtractive clones are capable of pinpointing chromosomal regions shared by N. meningitidis and N. gonorrhoeae which are involved in common aspects of the host interaction of both pathogens.

REFERENCE COUNT: 51 THERE ARE 51 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 17 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 1999:457991 BIOSIS
DOCUMENT NUMBER: PREV199900457991
TITLE: Meningococcal **vaccine**: Myth and reality.
AUTHOR(S): **Nassif, X. (1)**
CORPORATE SOURCE: (1) faculte Necker-Enfants-Malades, Inserm U 411, 156, rue de Vaugirard, 75015, Paris France
SOURCE: Archives de Pediatrie, (1999) Vol. 6, No. SUPPL. 3, pp. 647-649.
ISSN: 0929-693X.
DOCUMENT TYPE: Article
LANGUAGE: French

L20 ANSWER 18 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 8
ACCESSION NUMBER: 1998:71228 HCAPLUS
DOCUMENT NUMBER: 128:164910

THIS PAGE BLANK (USPTO)

TITLE: Genes and gene products specific to **pathogenicity** of **Neisseria** meningitidis, methods for obtaining them and their biological applications
INVENTOR(S): **Nassif, Xavier; Tinsley, Colin** ; Achtman, Mark; Ruelle, Jean-Louis; Vinals, Carla; Merker, Petra
PATENT ASSIGNEE(S): Institut National De La Sante Et De La Recherche Medicale (INSERM), Fr.; Max-Planck-Gesellschaft Zur Forderung Der Wissenschaften E.V., Berlin; Smithkline Beecham; Nassif, Xavier; Tinsley, Colin; Achtman, Mark; Ruelle, Jean-Louis; Vinals, Carla; Merker, Petra
SOURCE: PCT Int. Appl., 150 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: French
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9802547	A2	19980122	WO 1997-FR1295	19970711
WO 9802547	A3	19980409		

W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
FR 2751000	A1	19980116	FR 1996-8768	19960712
FR 2751000	B1	19981030		
AU 9736977	A1	19980209	AU 1997-36977	19970711
AU 730423	B2	20010308		
EP 951552	A2	19991027	EP 1997-933727	19970711

R: AT, BE, CH, DE, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
JP 2001504684	T2	20010410	JP 1998-505685	19970711
US 2002164603	A1	20021107	US 2001-928457	20010814

PRIORITY APPLN. INFO.:
FR 1996-8768 A 19960712
WO 1997-FR1295 W 19970711
US 1999-214759 B1 19990422

AB DNA sequences that are found in **Neisseria** meningitidis that are unique to it, specific to **pathogenesis**, and not found in N. gonorrhoeae, N. lactamica or N. cinerea are cloned by representational difference anal. A no. of genes assocd. with **pathogenesis** that are found in N. meningitidis and N.gonorrhoeae including the genes of biosynthesis of the polysaccharide capsule (frpA, frpC, porA), pilC, the genes for rotamase, IgA protease, pilin, transferring-binding proteins and opacity proteins and the sequence IS1106. The genes map in clusters in three regions of the chromosome. The gene products can be used as antigens in the raising of antibodies for diagnostic or therapeutic uses, e.g. specific immunoassays or **vaccines**. The roles of the genes in **pathogenesis** can be studied by

targeted deletion.

L20 ANSWER 19 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
DUPLICATE 9

ACCESSION NUMBER:	1998:377538 BIOSIS
DOCUMENT NUMBER:	PREV199800377538
TITLE:	Transferrin-binding proteins as candidates for a broadly cross-reactive vaccine against serogroup B meningococcal disease.
AUTHOR(S):	Quentin-Millet, M. J.; **Rokbi, B.**; Mignon, M.; Maitre-Wilmotte, G.; Danve, B.; **Renauld, G.**; Lissolo, L.
CORPORATE SOURCE:	Pasteur Merieux Connaught, 1541 avenue Marcel Merieux, 69280 Marcy l'Etoile Cedex France
SOURCE:	Biotecnologia Aplicada, (Jan.-March, 1998) Vol. 15, No. 1, pp. 41. Meeting Info.: Selected Papers from Congreso Biotechnologia Habana '97 Havana, Cuba December 1-6, 1997 ISSN: 0864-4551.
DOCUMENT TYPE:	Conference
LANGUAGE:	English

L20 ANSWER 20 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 10

ACCESSION NUMBER:	1997:491635 HCAPLUS
DOCUMENT NUMBER:	127:94113
TITLE:	Preparation of libraries of antibody genes in expression vectors capable of undergoing recombination for generation of new antibodies
INVENTOR(S):	Sodoyer, Regis; **Aujame, Luc**; Geoffroy, Frederique; **Bouchardon, Annabelle**
PATENT ASSIGNEE(S):	Pasteur Merieux Serums & Vaccins, Fr.; Sodoyer, Regis; Aujame, Luc; Geoffroy, Frederique; Bouchardon, Annabelle
SOURCE:	PCT Int. Appl., 37 pp. CODEN: PIXXD2
DOCUMENT TYPE:	Patent
LANGUAGE:	French
FAMILY ACC. NUM. COUNT:	1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9720923	A1	19970612	WO 1996-FR1938	19961204
W: CA, JP, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
FR 2741892	A1	19970606	FR 1995-14325	19951204
FR 2741892	B1	19980213		
CA 2239490	AA	19970612	CA 1996-2239490	19961204
EP 865487	A1	19980923	EP 1996-941701	19961204
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
US 6174708	B1	20010116	US 1999-101629	19990122
PRIORITY APPLN. INFO.:			FR 1995-14325 A	19951204
			WO 1996-FR1938 W	19961204

AB Libraries of genes for antibody light and heavy chains, including const. and variable regions, in bacteriophage or phagemid vectors

that are capable of undergoing in vivo recombination are described for use in the generation of new antibody types in bacterial hosts. One suitable mechanism is xis-mediated excision. Sep. vectors are used for light and heavy chain genes and the variable regions are manufd. as fusion proteins with a capsid protein to direct surface display. The vectors can undergo one or more irreversible recombinations to generate recombinant final vectors with possible new genes for variable regions. The method increases the efficiency of generation and detection of new antibody types.

L20 ANSWER 21 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
DUPLICATE 11
ACCESSION NUMBER: 1998:144177 BIOSIS
DOCUMENT NUMBER: PREV199800144177
TITLE: What do we know about the entry of Neisseria meningitidis into the meninges.
AUTHOR(S): Nassif, X. (1); Pujol, C.; **Tinsley, C.**; Morand, P.; Eugene, E.; Marceau, M.; **Perrin, A.**; Pron, B.; Taha, M.-K.
CORPORATE SOURCE: (1) INSERM U411, Lab. Microbiologie, Fac. Med. Necker-Enfants Malades, 156 Rue de Vaugirard, 75015 Paris France
SOURCE: Bulletin de l'Institut Pasteur, (Oct.-Dec., 1997) Vol. 95, No. 4, pp. 219-235.
ISSN: 0020-2452.
DOCUMENT TYPE: General Review
LANGUAGE: English

L20 ANSWER 22 OF 25 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1996:386028 HCAPLUS
DOCUMENT NUMBER: 125:56214
TITLE: Hemoglobin receptors from **Neisseriae** meningitidis and N. gonorrhoeae, cDNA sequences, and **vaccine** development
INVENTOR(S): Stojiljkovic, Igor; So, Magdalene; Hwa, Vivian; Heffron, Fred; **Nassif, Xavier**
PATENT ASSIGNEE(S): Oregon Health Sciences University, USA
SOURCE: PCT Int. Appl., 103 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9612020	A2	19960425	WO 1995-US13623	19951017
WO 9612020	A3	19960523		

W: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT
RW: KE, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5698438	A	19971216	US 1994-326670	19941018
US 6121037	A	20000919	US 1995-537361	19951002
AU 9540073	A1	19960506	AU 1995-40073	19951017

THIS PAGE BLANK (USPTO)

AU 705509	B2	19990527		
EP 793720	A2	19970910	EP 1995-938842	19951017
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE				
JP 10508469	T2	19980825	JP 1995-513482	19951017
NO 9701768	A	19970603	NO 1997-1768	19970417
FI 9701634	A	19970616	FI 1997-1634	19970417
US 6277382	B1	20010821	US 1997-817707	19970819
PRIORITY APPLN. INFO.:			US 1994-326670 A2	19941018
			US 1995-537361	19951002
			WO 1995-US13623 W	19951017

AB The present invention relates to novel bacterial Hb receptor proteins and genes that encode such proteins. The invention is directed toward the isolation, characterization, diagnostic and therapeutic use of bacterial Hb receptor proteins, nucleic acids encoding such proteins, recombinant expression constructs comprising such nucleic acids and cells transformed therewith, and antibodies and epitopes of such Hb receptor proteins. The invention relates particularly to Hb receptor proteins and genes encoding such proteins from **Neisseria** species, esp. N. meningitidis and serotypes thereof, and N. gonorrhoeae. Methods for the diagnostic and therapeutic use of the proteins, epitopes, antibodies and nucleic acids of the invention are also provided, including the use of the proteins, epitopes, antibodies and nucleic acids of the invention for the prodn. of **vaccines** effective in providing **immunization** of a human against infection by **pathogenic** bacteria of **Neisseria** species.

L20 ANSWER 23 OF 25 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 1997:275422 BIOSIS
DOCUMENT NUMBER: PREV199799574625
TITLE: Bacterial meningitis: Treatment and prevention strategies.
AUTHOR(S): Bingen, F.; Bourillon, A.; Clavaud, R.; Geslin, P.; Gicquel, B.; Guerin, N.; Livartowski, A.; **Nassif, X.**; Reinert, P.; Riou, J.-Y.; et al.
SOURCE: Bingen, E.; Bourillon, A.; Clavaud, R.; Geslin, P.; Gicquel, B.; Guerin, N.; Livartowski, A.; **Nassif, X.**; Reinert, P.; INSERM. (1996) pp. xi+167p. Bacterial meningitis: Treatment and prevention strategies. Meningites bacteriennes: Strategies de traitement et de prevention. Publisher: INSERM (Institut National de la Sante et de la Recherche Medicale) 101, rue de Tolbiac, 75654 Paris Cedex 13, France. ISBN: 2-85598-686-9.
DOCUMENT TYPE: Book
LANGUAGE: French

AB This monograph on treatment and prevention strategies for bacterial meningitis is based on approximately 500 articles and represents the collective work of a group of experts assembled by the INSERM, the French National Institute of Health and Medical Research. The book provides a useful reference to health care professionals, medical instructors and researchers, public health professionals, epidemiologists and those involved with health care policy. The first of the book's two sections represents a detailed analysis of the current world literature on various aspects of bacterial meningitis that was developed during a series of meetings of the

expert group. The first part of this analytical section is devoted to general aspects of bacterial meningitis and contains chapters on clinical data and pathophysiology, germs and diagnostic methods, epidemiological and socioeconomic data. The following part discusses treatment and **vaccinal** prevention of bacterial meningitis, including antibiotic treatment and prophylaxis, current **vaccines** and **vaccine** strategies for dealing with four of the bacterial meningitis **pathogens** (Haemophilus influenzae- b, **Neisseria** meningitidis, Streptococcus pneumoniae and Mycobacterium tuberculosis). The third part of the analytical section focuses on prospects for **vaccines** against meningitis, especially **vaccines** against the **Neisseria** meningitidis, Streptococcus pneumoniae and Mycobacterium tuberculosis **pathogens**. The second section of the book provides a synthesis of the information available on bacterial meningitis, its treatment and the research and development of **vaccines** for its prevention. This section concludes with seven recommendations developed by the group of experts for the therapeutic management and prevention of bacterial meningitis. Tables, graphs, diagrams and end-of-chapter bibliographies supplement the text.

L20 ANSWER 24 OF 25 MEDLINE

ACCESSION NUMBER: 96009214 MEDLINE
DOCUMENT NUMBER: 96009214 PubMed ID: 7553571
TITLE: Interaction of **pathogenic** **neisseriae** with nonphagocytic cells.
AUTHOR: **Nassif X**; So M
CORPORATE SOURCE: Institut National de la Sante et de la Recherche Medicale U411, Faculte de Medecine Necker-Enfants Malades, Universite Rene Descartes, Paris, France.
SOURCE: CLINICAL MICROBIOLOGY REVIEWS, (1995 Jul) 8 (3) 376-88. Ref: 148
Journal code: 8807282. ISSN: 0893-8512.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199510
ENTRY DATE: Entered STN: 19951227
Last Updated on STN: 19951227
Entered Medline: 19951025

AB The ability to interact with nonphagocytic cells is a crucial virulence attribute of the meningococcus and the genococcus. Like most bacterial **pathogens, Neisseria** meningitidis and **Neisseria** gonorrhoeae initiate infections by colonizing the mucosal epithelium, which serves as the site of entry. After this step, both bacteria cross the intact mucosal barrier. While N. gonorrhoeae is likely to remain in the subepithelial matrix, where it initiates an intense inflammatory reaction, N. meningitidis enters the bloodstream, and eventually the cerebrospinal fluid to cause meningitis. Both **pathogens** have evolved very similar mechanisms for interacting with host cells. Surface structures that influence bacterium-host interactions include pili, the meningococcal class 5 outer membrane proteins or the gonococcal opacity proteins, lipooligosaccharide, and the

meningococcal capsule. This review examines what is known about the roles these structures play in bacterial adhesion and invasion, with special emphasis, on pilus-mediated adhesion. Finally, the importance of these structures in **neisserial pathogenesis** is discussed.

L20 ANSWER 25 OF 25 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 12
ACCESSION NUMBER: 1993:20647 HCAPLUS
DOCUMENT NUMBER: 118:20647
TITLE: Antibodies recognizing a variety of different structural motifs on meningococcal Lip antigen fail to demonstrate bactericidal activity
AUTHOR(S): **Tinsley, Colin R.**; Virji, Mumtaz; Heckels, John E.
CORPORATE SOURCE: Med. Sch., Southampton Univ., Southampton, SO9 4XY, UK
SOURCE: Journal of General Microbiology (1992), 138(11), 2321-8
CODEN: JGMIAN; ISSN: 0022-1287
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The **neisserial** Lip antigen is a conserved antigen assocd. with the **pathogenic Neisseria** species, and is composed of multiple repeats of a consensus pentapeptide. A series of monoclonal antibodies reacting with meningococcal Lip antigen were subjected to epitope mapping, using solid-phase synthetic peptides based on the concensus repeat sequence. The antibodies were found to recognize different continuous epitopes based on the consensus sequence. One monoclonal antibody was utilized in affinity chromatog. to obtain purified Lip antigen and the antigen was used for **immunization** of mice. The resulting antisera did not recognize Lip antigen on Western blots but reacted specifically with Lip antigen in immune pptn. expt., indicating that the predominant polyclonal immune response was directed against conformational epitopes. Despite the diversity of both continuous and conformational epitopes recognized by the antibodies produced, none of the antibodies demonstrated the ability to promote complement-mediated bactericidal activity. Thus, despite its initial apparent promise as a potential **vaccine** candidate the case for the inclusion of Lip antigen in **vaccine** formulation cannot be supported at present.

FILE 'HOME' ENTERED AT 10:19:09 ON 28 JAN 2003